



**Universidad Nacional Mayor de San Marcos**

**Universidad del Perú. Decana de América**

Facultad de Ingeniería de Sistemas e Informática  
Escuela Académico Profesional de Ingeniería de Sistemas

**Data mart para el análisis post-operatorio de cáncer de  
páncreas en el Instituto Nacional de Enfermedades  
Neoplásicas**

**TESINA**

Para optar el Título Profesional de Ingeniero de Sistemas

**AUTOR**

Pedro Manuel ZAVALA QUIJANO

**ASESOR**

Virginia VERA POMALAZA

Lima, Perú

2008



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

## Referencia bibliográfica

---

Zavala, P. (2008). *Data mart para el análisis post-operatorio de cáncer de páncreas en el Instituto Nacional de Enfermedades Neoplásicas*. Tesina para optar el título profesional de Ingeniero de Sistemas. Escuela Académico Profesional de Ingeniería de Sistemas, Facultad de Ingeniería de Sistemas e Informática, Universidad Nacional Mayor de San Marcos, Lima, Perú.

---

Este trabajo está dedicado a toda mi familia y en especial a mi madre por su apoyo todos estos años, por su infinito amor, comprensión y por ayudarme a que este momento llegara.

## AGRADECIMIENTOS

A la profesora Virginia Vera Pomalaza, por su orientación y dedicación para que éste trabajo cumpla con los objetivos trazados.

A los profesores de la UNMSM, por sus observaciones teóricas que me sirvieron de mucho.

Al Médico Oncólogo Hugo Valencia, por sus consejos, observaciones y apoyo incondicional en la elaboración de este trabajo.

A todas personas que indirectamente me ayudaron para culminar éste trabajo y que muchas veces constituyen un invalorable apoyo.

Y por encima de todo doy gracias a Dios.

# **DATAMART PARA EL ANÁLISIS POST-OPERATORIO DE CÁNCER DE PÁNCREAS EN EL INSTITUTO NACIONAL DE ENFERMEDADES NEOPLÁSICAS**

## **RESUMEN**

El presente trabajo, propone un método para la elaboración de un sistema de ayuda a la toma de decisiones a fin de poder realizar análisis post – operatorios para el cáncer de páncreas en el Departamento de Abdomen del Instituto Nacional de Enfermedades Neoplásicas.

Aquí se complementan técnicas de Data mart, de extracción y carga de datos, para poder realizar la explotación de la información y poder tomar decisiones y gestionar los indicadores de calidad propios de la institución.

Palabras Claves: Data mart, cáncer de páncreas, Inteligencia de Negocios.

# **DATAMART FOR THE POSTOPERATORY ANALYSIS OF PANCREATIC CANCER AT THE NATIONAL NEOPLASICS DISEASES INSTITUTE**

## **ABSTRACT**

The present work propose a method for the elaboration of a system of help to the make of decisions in order to be able to realize analyses post - operative for the cancer of pancreas in the National Neoplastic Diseases Institute.

Here it complements each other techniques of data mart, of extraction and load of information, to be able to realize the exploitation of the information and be able to take decisions and manage the indicators of quality own of the institution.

Key Words: Data mart, pancreatic cancer, Business Intelligence.

## ÍNDICE DE FIGURAS

Figura 2.1: Partes del Páncreas.....	13
Figura 2.2: El ciclo de Business Intelligence .....	20
Figura 2.3: Las primeras etapas evolutivas del manejo de información.....	27
Figura 2.4: La naturaleza del proceso de Extraer.....	28
Figura 2.5: Falta de credibilidad en la arquitectura de crecimiento natural.....	29
Figura 2.6: El Ciclo de Vida de Desarrollo de Sistemas para el ambiente DW.....	30
Figura 2.7: Ejemplos de Transformación.....	33
Figura 2.8: Cubo de datos.....	34
Figura 2.9: Ocurrencias en la tabla de hechos.....	34
Figura 2.10: Acumulados según dos y las tres dimensiones.....	35
Figura 2.11: Acumulados según todas las dimensiones.....	35
Figura 2.12: Unión del cubo y sus acumulados.....	36
Figura 2.13: El cubo de medidas con todos sus acumulados .....	36
Figura 2.14: Modelo estrella para procesar órdenes.....	37
Figura 2.15: Diferencias entre el modelo dimensional y el modelo relacional.....	38
 Figura 3.1: Diagrama Funcional (Star Net).....	 52
Figura 3.2: Esquema de Data Mart para un Consultorio.....	53
Figura 3.3: Pasos de la Metodología.....	54
 Figura 4.1: Esquema del Data Mart .....	 57
Figura 4.2: Ejemplo de dimensiones usadas en cubo .....	59
Figura 4.3: Tabla de Hecho.....	61
Figura 4.4: Cubo 1 del prototipo.....	63
Figura 4.5: Cubo 2 del prototipo .....	64
Figura 4.6: Cubo 3 del prototipo.....	65
 Figura 5.1: Desarrollo de la BD en SQL Server 2005 para el uso del Data mart .....	 68
Figura 5.2: Diagrama de la Base de Datos en MSSQL Server Management Studio.....	69
Figura 5.3: Desarrollo de la solución BI en Business Intelligence Development Studio.	70
 Figura 6.1: Resultados de la Elaboración del Cubo1.....	 71
Figura 6.2: Representación gráfica del Cubo 1.....	62
Figura 6.3: Representación gráfica del Cubo 2.....	73
Figura 6.4: Representación gráfica de resultados del Cubo 2.....	74
Figura 6.5: Representación gráfica del Cubo 3.....	75



# ÍNDICE

RESUMEN.....	v
ABSTRACT.....	vi
ÍNDICE DE FIGURAS.....	vii
ÍNDICE.....	1
INTRODUCCIÓN.....	3
 CAPÍTULO 1: PLANTEAMIENTO DEL PROBLEMA.....	 4
1.1. ANTECEDENTES.....	4
1.2. DEFINICIÓN DEL PROBLEMA.....	6
1.3. OBJETIVOS.....	6
1.4. JUSTIFICACIÓN.....	6
1.5. PROPUESTA.....	7
 CAPÍTULO 2: MARCO TEÓRICO.....	 10
2.1 CÁNCER.....	10
2.2 CÁNCER DE PÁNCREAS.....	11
2.2.1. Introducción.....	11
2.2.2. Tumores del páncreas exocrino.....	12
2.2.3. Tumores Quísticos(TQ) .....	15
2.2.4. Tumores Ampulares y Periapulares.....	16
2.3 INTELIGENCIA DE NEGOCIOS.....	17
2.3.1 Inteligencia de Negocios (Business Intelligence) .....	17
2.3.2 Data Warehouse.....	31
2.3.3 Migración de Datos: Extracción, Transformación y Carga.....	31
2.3.4 Data Mart.....	33
2.3.5 Bases De Datos Médica.....	34
2.3.6 Cubo de Información.....	34
2.3.7 Modelo Dimensional.....	37
2.3.8 Definición de Términos Básicos.....	39
 CAPÍTULO 3: ESTADO DEL ARTE.....	 42
3.1. DISTRIBUCIÓN A NIVEL NACIONAL.....	42
3.2. TÉCNICAS Y MÉTODOS.....	43
3.3. CASO DE ESTUDIO .....	53
CASO 1: DATAMART PARA UN CONSULTORIO.....	53
CASO 2: UNA METODOLOGÍA PARA SECTORIZAR PACIENTES.....	54
3.3.1. ESTUDIO DE CASOS.....	55
 CAPÍTULO 4: RESOLUCIÓN DEL PROBLEMA APLICANDO LA TÉCNICA SELECCIONADA.....	 56

4.1 PROPUESTA METODOLÓGICA .....	56
4.2 IMPLEMENTACIÓN.....	57
4.2.1 Diseñar el Esquema Dimensional del Data Mart.....	57
4.3 TABLA DE HECHO .....	60
4.4 DEFINICIÓN DE INDICADORES HOSPITALARIOS.....	61
4.5 INDICADORES DE CALIDAD.....	61
4.5.1 Porcentaje de Morbilidad.....	61
4.5.2 Porcentaje de Mortalidad.....	61
4.5.3 Porcentaje de Mortalidad Post-Operatoria.....	61
4.6 CUBOS DE INFORMACIÓN.....	62
Cubo 1: .....	62
Cubo 2: .....	63
Cubo 3: .....	64
CAPÍTULO 5: DESCRIPCIÓN DE LA SOLUCIÓN TECNOLÓGICA.....	65
5.1 HERRAMIENTAS UTILIZADAS.....	65
5.1.1 Introducción.....	65
5.1.2 ETL.....	65
5.1.3 Microsoft SQL Server Management Studio 2005.....	66
5.1.4 Business Intelligence Development Studio.....	67
5.1.5 Informes del Data Mart.....	69
CAPÍTULO 6: ANÁLISIS E INTERPRETACIÓN DE RESULTADOS.....	70
CAPÍTULO 7: CONCLUSIONES Y FUTUROS TRABAJOS.....	76
ANEXOS.....	78
TABLA 1: .....	78
TABLA 2: .....	79
TABLA 3: .....	80
REFERENCIAS BIBLIOGRÁFICAS.....	81

## **INTRODUCCIÓN:**

La cirugía del páncreas ha evolucionado favorablemente en los últimos años, disminuyendo significativamente la mortalidad operatoria; sin embargo, el pronóstico de los pacientes con cáncer de páncreas continúa siendo sombrío y es recién con el advenimiento de la tecnología moderna se puede realizar toma de decisiones con mayor prontitud.

Este trabajo se divide en 4 partes, cada una de las cuales se describe a continuación, a manera resumida:

En el Capítulo I se realiza una explicación de los Antecedentes del problema, la definición del problema, la justificación, y se define los objetivos generales y específicos de la tesis. Por último se define la Propuesta del trabajo.

En el Capítulo II se define el Marco teórico, donde se fundamenta las actuales situaciones por las que pasa toda institución para poder obtener conocimiento. Se empieza definiendo el antecedente, los conceptos con los que se ha contado en el proyecto, para poder irse introduciendo en el trabajo realizado. Tratamos sobre temas como cáncer de páncreas, Data Warehouse (definición y arquitectura tanto física como lógica), Data Mart, Almacenamiento OLAP (cubos, dimensiones, métricas, jerarquías y estrategias de almacenamiento). Veremos también, conceptos sobre transformación y carga de datos.

En el Capítulo III se realiza una visión general del proyecto, se define y explica los pasos que comprende la metodología, y se desarrollan los mismos. Aquí es donde doy a conocer la fuente de datos con los que contaremos en el proyecto, vemos como diseñamos y poblamos nuestro Data Mart.

En el Capítulo IV se hace una breve explicación del instituto en estudio, para ubicar al lector en el contexto deseado. Se detallan los resultados, discusión e interpretación de la investigación. Finalmente se presentan conclusiones y se plantean algunas recomendaciones que se pueden aplicar, que complementen el trabajo presentado.

## **CAPÍTULO 1: PLANTEAMIENTO DEL PROBLEMA**

### **1.1. Antecedentes**

El cáncer de páncreas es la segunda neoplasia maligna del tracto gastrointestinal en orden de frecuencia. También es el de peor pronóstico; la sobrevivida al año es de 20%, e inferior a 5% a los cinco años. El único tratamiento que puede ser curativo es la extirpación quirúrgica. Los tratamientos alternativos se reservan para los casos de enfermedad más extendida, pero solo mitigan temporariamente los síntomas [Anderson06].

Si bien la mortalidad operatoria de la pancreatectomía<sup>1</sup> ha descendido hasta el 4% en centros con experiencia, la morbilidad puede llegar a ser del 50%. Entre las causas de esta morbilidad destacan el vaciamiento gástrico retrasado, la fístula pancreática, absceso abdominal, hemorragia, infección de la herida, distress respiratorio, hemorragia post operatoria intra-abdominal, etc.

Se consideró este problema un tema clave para el cumplimiento de indicadores de calidad establecidos en el Instituto Nacional de Enfermedades Neoplásicas (INEN) como parte de su política institucional, sin embargo, cabe indicar que hasta el presente no se han desarrollado trabajos sobre Data Mart, Datamining y Toma de decisiones en el ámbito del análisis pre-operatorios o post-operatorios y específicamente en el INEN.

Asimismo, con relación a las variables del tema, no se han encontrado investigaciones que hayan abordado estos temas aplicados a la problemática

---

<sup>1</sup> Cirugía para extraer todo el páncreas o una parte del mismo. [Dorland97]

planteada, con lo cual consideramos que la presente investigación reúne las condiciones metodológicas suficientes para ser considerada inédita.

## **1.2. Definición del Problema**

Las organizaciones dedicadas a la atención de la salud, como muchas de sus pares en otras áreas, asisten a un proceso de creciente informatización. La mayor parte de las aplicaciones aún se vinculan con procesos netamente administrativo-contables, pero el grado de informatización de datos estrictamente médicos es cada vez mayor. Las Base de Datos Transaccionales propias de la organización médica en estudio no escapa a los problemas que afectan a las organizaciones de los otros sectores, y los analistas se enfrentan a los mismos problemas de “encarcelamiento” de los datos.

El Instituto Nacional de Enfermedades Neoplásicas cuenta con una gran cantidad de información contenida en su Sistema Integrado Hospitalario, ésta solo registra información básica de las historias clínicas, informes de laboratorio y de algunos estudios anatómo-patológicos, no se la considera como lo que es realmente; es decir, una fuente importante de conocimiento que puede ser útil para la institución.

En entrevistas con los médicos oncólogos del INEN, se supo que no existe suficiente información actualizada para el análisis post-operatorio, que con un buen criterio se impediría intervenciones quirúrgicas innecesarias. Además surge la necesidad de obtener información basándose en nuevos y diversos criterios, para lo que el médico oncólogo al no contar con algún medio de ayuda a la toma de decisiones y ante la ausencia del uso de Inteligencia de Negocios en la institución, los médicos toman decisiones, analizan riesgos y oportunidades basados en información personal o anecdótica, incompleta, desactualizada o en la experiencia del médico.

A esto se suma la falta de estandarización en el almacenamiento digital de las historias clínicas para mejorar el tiempo de respuesta en la búsqueda de la información del paciente y diagnósticos realizados. Éste es uno de los problemas con más repercusión en los procesos del hospital en estudio.

### 1.3. Objetivos:

Objetivo General:

- Definir una metodología para el modelado y construcción de un Data Mart, que ayude a la toma de decisiones para el análisis post-operatorio de cáncer de páncreas en el Instituto Nacional de Enfermedades Neoplásicas.

Y teniendo como objetivos secundarios:

- Demostrar como la transformación de la información correspondiente a un período de trabajo del sistema actual almacenado en tablas planas hacia una Base de Datos Relacional permite establecer objetivos a alcanzar en la toma de decisiones.
- Conocer como el modelar, construir y cargar una base de datos relacional a un Data mart coadyuva a establecer estrategias para la Toma de Decisiones.
- Usar el Modelamiento Dimensional como base de los cubos de información según los campos requeridos en la Tabla de Hechos.
- Conocer las fases del extracción de datos de bases de datos y la importancia de las mismas en el éxito del proceso (en especial las de limpieza y selección de datos), lo que se conoce como *Extract, Transform and Load* (ETL).
- Mostrar correctamente la información clasificada en base a los cubos de información, para su posterior visualización en reportes para el usuario.

### 1.4. Justificación

La justificación del trabajo, se centra en el hecho de que se propone una solución al requerimiento de un problema que actualmente tiene el Departamento de Abdomen del Instituto Nacional de Enfermedades Neoplásicas (INEN).

La no existencia de una herramienta de toma de decisiones para nos análisis de control post-operatorio en pacientes con cáncer de páncreas, hace que no se

puedan medir oportunamente el diagnóstico post-operatorio así como los índices indicadores de calidad establecidos por la institución.

En fin, se convierte en un aporte importante, con futuras mejoras, que podrá servir como fuente de conocimiento no sólo para el Departamento de Abdomen, sino también para todos los Departamentos del Instituto Nacional de Enfermedades Neoplásicas.

### **1.5. Propuesta:**

La propuesta radica en el desarrollo de una solución ante el problema citado debido a la difícil tarea de medir los indicadores de calidad para evaluar los estados post-operatorios de todos los pacientes que han sido operados de cáncer de páncreas.

Al lograr un sistema de ayuda orientado a la toma de decisiones se lograría una sustancial mejora para el análisis de los indicadores de calidad.

Con el desarrollo de la solución se estaría agilizando la entrega de reportes, análisis de los resultados para hacer un seguimiento tras la cirugía y tomar acciones correctivas sobre la mortalidad, morbilidad operatoria del cáncer de páncreas y demás indicadores para una mejor gestión hospitalaria.

## **CAPÍTULO 2: MARCO TEÓRICO**

### **2.1 CÁNCER:**

Entre las definiciones de cáncer tenemos:

- i. Tejido celular en el que los mecanismos que controlan el crecimiento normal están permanentemente alterados, permitiendo su crecimiento progresivo [Solidoro+75].
- ii. Un grupo de enfermedades de causa desconocida, probablemente múltiple que ocurre en todas las poblaciones humanas y animales, apareciendo en todos los tejidos compuestos por células con potencialidad de división [Solidoro05].
- iii. Una enfermedad genética de las células somáticas, que resulta de la mutación de una célula previamente normal, en una nueva célula permanentemente maligna [Solidoro+75].
- iv. Una enfermedad de organización tisular, donde no existe “célula cancerosa” como tal, sino células que se comportan de una manera que es definida arbitrariamente como cancerosa [Solidoro+75].
- v. Una enfermedad espontáneamente autoagresiva de un tejido iniciada por una mutación genética, ocurrida por azar en una célula madre [Solidoro+75].

### **Hechos Básicos:**

Independientemente a los intentos de definir un grupo de hechos básicos debe ser tenido en cuenta:

- Las neoplasias pueden originarse sólo de células con capacidad de proliferar.
- Las células tumorales no se ‘diferencian’, pueden madurar en forma inadecuada llamándose “anaplásicas” o “indiferenciadas”.



- Los cánceres pueden surgir como consecuencia de una variedad de estímulos (químicos, físicos o virales), requiriéndose por lo general, un prolongado periodo de latencia.
- Las neoplasias son frecuentemente multicéntricas dentro del mismo tejido, por ejemplo el aparente aumento en la incidencia de cáncer de mama bilateral y bifocal.
- La hiperplasia y displasia preceden a menudo el desarrollo de muchos tumores en meses y años (el reconocimiento y la terapia apropiada en éste estado, pueden prevenir muchos cánceres).
- Las células cancerosas pueden permanecer latentes por periodos prolongados.
- Pocas neoplasias que regresan espontáneamente. El comportamiento biológico de otros, pueden ser modificados por inmunología, hormonas y quimioterapia.

El diagnóstico y la definición de cáncer casi siempre dependen del examen microscópico; sin embargo, es de gran importancia recordar que la histología del cáncer, no es equivalente a la clínica del cáncer o al estadio lesional observado en un determinado paciente.

## **INFORMACIÓN SOBRE LOS ESTADÍOS**

El sistema de clasificación para el cáncer pancreático exocrino todavía se encuentra en proceso de evolución. La importancia del estadio, más allá de si es resecable o no resecable, es incierta, ya que los tratamientos más avanzados han demostrado tener poco efecto en la supervivencia. Sin embargo, para poder contar con una definición uniforme de la enfermedad, es necesario tener conocimiento de su extensión. El cáncer del páncreas generalmente se identifica según el sitio afectado dentro del páncreas. El enfoque quirúrgico difiere para los tumores situados en la cabeza, el cuerpo, la cola o la apófisis unciforme del páncreas [Polanco99].

El Comité Estadounidense Conjunto sobre el Cáncer (AJCC, por sus siglas en inglés) ha designado los estadios mediante clasificación TNM [Solidoro05].

## **Definiciones TNM**

### **Tumor primario (T)**

- TX: El tumor primario no puede evaluarse.
- T0: No hay prueba de tumor primario.
- Tis: Carcinoma in situ.
- T1: Tumor está limitado al páncreas y mide de 2 cm o menos en su diámetro mayor.
- T2: Tumor está limitado al páncreas y mide de 2 cm o más en su dimensión mayor.
- T3: El tumor se extiende más allá del páncreas pero sin implicación alguna del tronco celíaco o la arteria mesentérica superior.
- T4: El tumor comprende el tronco celíaco o la arteria mesentérica superior (tumor primario irresecable).

### **Ganglios linfáticos regionales (N)**

- NX: Los ganglios linfáticos regionales no pueden evaluarse.
- N0: No hay metástasis a los ganglios linfáticos regionales.
- N1: Existe metástasis a los ganglios linfáticos regionales.

### **Metástasis a distancia (M)**

- MX: La metástasis a distancia no puede evaluarse.
- M0: No hay metástasis a distancia.
- M1: Existe metástasis a distancia.

Las definiciones TNM son utilizadas para definir los Grupos de Etapas del Cáncer de Páncreas, como se muestra en la Tabla 2.1

<b>GRUPOS DE ETAPAS DEL CÁNCER DE PÁNCREAS SEGÚN LA AJCC.</b>			
<b>Etapa</b>	<b>T</b>	<b>N</b>	<b>M</b>
<b>Etapa 0</b>	<b>Tis</b>	<b>N0</b>	<b>M0</b>
<b>Etapa I</b>	<b>T1</b>	<b>N0</b>	<b>M0</b>
	<b>T2</b>	<b>N0</b>	<b>M0</b>
<b>Etapa II</b>	<b>T3</b>	<b>N0</b>	<b>M0</b>
<b>Etapa III</b>	<b>T1</b>	<b>N1</b>	<b>M0</b>
	<b>T2</b>	<b>N1</b>	<b>M0</b>
	<b>T3</b>	<b>N1</b>	<b>M0</b>
<b>Etapa IV-A</b>	<b>T4</b>	<b>Cualquier N</b>	<b>M0</b>
<b>Etapa IV-B</b>	<b>Cualquier T</b>	<b>Cualquier N</b>	<b>M1</b>

Tabla 2.1: Grupo de Estadíos Clínico – Patológico [Solidoro+75].

## 2.2 CÁNCER DE PÁNCREAS.

### 2.2.1. INTRODUCCIÓN

Se describirá la patología tumoral pancreática de acuerdo a los siguientes capítulos [Nago94]:

2.2.2 Tumores del Páncreas Exocrino.

2.2.3 Tumores Quísticos.

2.2.4 Tumores Ampulares y Periapulares.

Al ser los tumores del páncreas endocrino sumamente infrecuentes no serán incluidos. Por otro lado, la inclusión del tercer grupo se hace en consideración al parecido en el comportamiento clínico con los tumores malignos de la cabeza del páncreas, aunque con distinto grado de evolución y pronóstico y cuyo tratamiento quirúrgico es prácticamente el mismo.

### 2.2.2. TUMORES DEL PÁNCREAS EXOCRINO

El páncreas exocrino, formado por acinos y conductos excretores, puede generar tumores a partir de estas estructuras. Estos tumores pueden ser benignos o malignos.

- **Tumores Benignos:**

- Adenoma acinar (muy infrecuente).

- **Tumores Malignos.**

- Carcinoma acinar (apenas el 2%).
- Adenocarcinoma ductal (muy frecuente).

Tanto los Adenomas como el Carcinoma acinar son de rara presentación por ello describiremos únicamente el Adenocarcinoma ductal conocido comúnmente como Cáncer de Páncreas.

### CÁNCER DE PÁNCREAS (CP)

Se origina en el epitelio ductal y es el tumor maligno más frecuente del páncreas (90%). Se lo asocia con tabaquismo, anomalías ductales (páncreas divisum) y pancreatitis crónica alcohólica y más si ésta es calcificante. Más frecuente después de los 60 años, con predominio en hombres en una relación de 3 -1 con respecto a las mujeres. Su etiología permanece aún desconocida [Estremadoyro+96].

En los Estados Unidos su incidencia es de 9.1 por cada cien mil habitantes y es la cuarta causa de muerte por cáncer. En nuestro país, información del Centro de Investigación de Cáncer Maes Heller de Lima Metropolitana (1990-1993) refiere que sobre un total de 28,668 casos de cáncer, el Cáncer de Páncreas ocupa el décimo séptimo lugar con 553 casos, que equivale a una incidencia de 3,9 por cien mil habitantes, ocupando el cuarto lugar de las neoplasias malignas del aparato digestivo, precedido por el Cáncer de estómago, colon y recto. [RCLM98].

Esquemáticamente podemos decir que el 70% se desarrolla en la cabeza, el 20% en el cuerpo y 10% en la cola del páncreas. Las partes del páncreas se muestran en la figura 2.1.

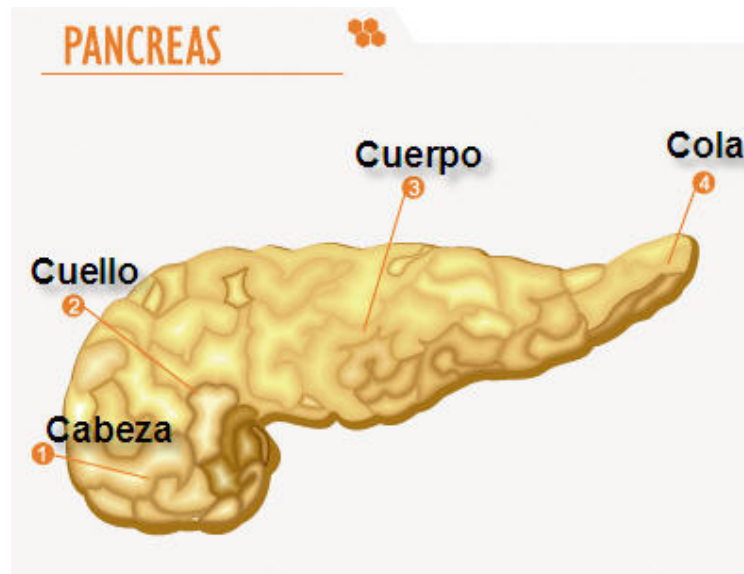


Figura 2.1: Partes del Páncreas.

- **Cáncer de Cabeza de Páncreas**

Macroscópicamente se presenta como un agrandamiento de la porción afectada. Tumor de tamaño pequeño de consistencia dura, mal delimitado y con una reacción peritumoral (pancreatitis paraneoplásica) que magnifica el tamaño del tumor.

El cáncer cefálico por lo general obstruye el conducto de Wirsung originando una dilatación retrógrada de 3 cm o más de diámetro (siendo el diámetro normal de 2 mm).

El colédoco retropancreático también es invadido, produciendo una obstrucción total, con dilatación de la vía biliar intra y extrahepática. La vesícula de paredes no inflamadas con cístico permeable también se dilata. La infiltración compromete rápidamente vasos mesentéricos superiores, vena porta y órganos vecinos (duodeno estómago) y estructuras ganglionares retro-peritoneales originando en gran porcentaje de casos su incurabilidad.

- **Cáncer de Cuerpo de Páncreas**

A las características macroscópicas ya mencionadas se suma una invasión precoz de los nervios espláncnicos, pudiendo invadir más tardíamente la vía biliar por avance de la neoplasia o metástasis ganglionares.

- **Cáncer de Cola de Páncreas**

En esta localización, tiende a desarrollar formaciones tumorales de regular volumen con invasión de órganos vecinos y estructuras retroperitoneales que pueden ser palpadas al examen clínico.

Desde el punto de vista histológico estos adenocarcinomas son de distinto grado de diferenciación.

Existen otros tipos histológicos menos frecuentes de origen ductal: adenoescamosos (acantomas), de células claras, de células en anillo de sello y los carcinomas anaplásicos. Estos últimos conocidos también con el nombre de pleomorfos o sarcomatoides corresponden al 7% de los tumores pancreáticos y se ubican de preferencia en el cuerpo o cola del páncreas.

- **Diseminación**

El Cáncer de Páncreas se expande por invasión directa a órganos vecinos. Por vía linfática a ganglios vecinos, regionales y ganglios alejados supra e inframesocolónicos. Por vía venosa da metástasis en hígado y finalmente metástasis generalizadas.

### **Cáncer de cabeza de Páncreas**

Al inicio tiene síntomas inespecíficos, de carácter general: trastornos dispépticos, astenia, anorexia y pérdida de peso (Síndrome general).

La ictericia está presente prácticamente en el 100% de los casos. Como se mencionó, el Cáncer de Páncreas invade rápidamente el colédoco originando una

ictericia obstructiva con dilatación de la vía biliar intra y extrahepática de carácter progresivo y con bilirrubina a predominio directa (Síndrome icterico). El paciente refiere además coluria y acolia.

Pueden aparecer otros síntomas inconstantes como náuseas, vómitos y diarrea. Los primeros cuando ha sido invadido el duodeno y la diarrea como expresión de síndrome de mala absorción.

### **Cáncer de Cuerpo y Cola:**

En el Cáncer de Cuerpo el síntoma predominante es el dolor. Es de gran intensidad, de localización epigástrica con irradiación a hipocondrio izquierdo y región lumbar del mismo lado, es continuo y con exacerbaciones postprandiales, con el reposo nocturno y en decúbito dorsal. El paciente adopta posiciones antálgicas de cuclillas y el tórax inclinado hacia adelante. Se interpreta este síndrome doloroso como la invasión del plexo solar por la neoplasia. La pérdida de peso y la anorexia, asociado al dolor, completa la tríada diagnóstica. La ictericia no es muy frecuente y cuando aparece lo hace tardíamente por invasión tumoral y más frecuentemente por metástasis ganglionares.

En el Cáncer de Cola, al síndrome general, se le agrega la tumoración palpable. Al examen físico encontramos: tumor palpable, frecuente en Cáncer de Cola e inconstante en el de Cuerpo; hepatomegalia y ascitis por metástasis generalizadas.

### **2.2.3. TUMORES QUÍSTICOS (TQ)**

Son tumores raros, apenas el 3% de los tumores pancreáticos y sólo el 10% son malignos.

Desde el punto de vista anatomopatológico los TQ se dividen en:

- 1- Quistes Verdaderos: Con revestimiento epitelial.
- 2- Pseudoquistes: Carecen de revestimiento. Forman parte de las complicaciones de las Pancreatitis.

#### **2.2.4. TUMORES AMPULARES Y PERIAMPULARES**

Se denomina así a los tumores malignos que se localizan en la desembocadura común del colédoco y del Wirsung en el duodeno o en su vecindad. Éstos tumores se originan en el epitelio mucoso del duodeno, de la Ampolla de Vater, del Colédoco Inferior, del Wirsung terminal o Islotes Pancreáticos Cefálicos que rodean a las estructuras mencionadas. Algunos autores no consideran los dos últimos por ser neoplasias genuinas del páncreas. Por tener características histológicas y clínicas semejantes justifican tratarlos en conjunto.

Con excepción de los tumores ampulares, en los periampulares es difícil apreciar macroscópicamente su origen. La histología es similar (adenocarcinomas) con distinto grado de diferenciación.

Debemos enfatizar que los tumores ampulares y periampulares oportunamente operados pueden ser curables dado su bajo grado de malignidad, en contraposición al Cáncer de Páncreas.



## 2.3 INTELIGENCIA DE NEGOCIOS

En esta parte se explican las Bases usadas para el desarrollo de este trabajo, detallando que es una Herramienta Inteligente (*Business Intelligence*), entre otros.

### 2.3.1 Inteligencia de Negocios (*Business Intelligence*)<sup>2</sup>

El término *Business Intelligence*, también conocido como BI, es usado por diferentes expertos y fabricantes de Software para distinguir un amplio rango de tecnologías, plataformas de software, aplicaciones específicas y procesos. Podemos verlo desde 3 perspectivas:

#### Tomar mejores decisiones rápidamente

El objetivo primario del *Business Intelligence* es ayudar a la gente a tomar decisiones que mejoren el rendimiento de la compañía e impulsen su ventaja competitiva en el mercado, en forma breve, *Business Intelligence* faculta a las organizaciones a tomar las mejores decisiones rápidamente.

*Business Intelligence* ayuda a tomar mejores decisiones analizando si las acciones tomadas están de hecho dando resultados hacia los objetivos de la compañía. Decidir cuál es la mejor decisión para una organización, se lleva mejor a cabo con un definido y claro conjunto de objetivos y un plan para conseguirlos.

La necesidad de ser veloz también se aplica a la ganancia de reacción dentro de una organización. Si un acceso y procesamiento rápido de la información no son posibles, las decisiones son tomadas sin información o con información antigua o discontinuada. Las consecuencias negativas de contar con información o con información discontinuada pueden ser enormes, tal como perder un cliente clave o continuar produciendo un producto que los clientes no desean.

Aunque disponer de información perfecta para apoyar todas las tomas de decisiones es un objetivo inalcanzable, no hay dudas de que tomar de forma continua decisiones adecuadas en un plazo corto proporcionarían una ventaja competitiva.

---

<sup>2</sup> De "Business Intelligence : Técnicas de Análisis para la toma de decisiones estratégicas " Mc Graw Hill 2003.

### Convertir los datos en información

Para tomar mejores decisiones más rápidamente, los directivos y gerentes necesitan de información relevante y útil al alcance de la mano. Pero es común una larga brecha entre la información que los responsables en la toma de decisiones requieren y las grandes cantidades de datos que las organizaciones recopilan cada día. Nosotros llamamos a esto la brecha de análisis (análisis gap).

Para saltar esta brecha, las organizaciones hacen significativas inversiones en desarrollar sistemas de BI para convertir los datos originales o “crudos” en información de utilidad. Los sistemas de BI más efectivos tienen acceso a inmensas cantidades de datos (medidas en gigabytes incluso en terabytes) para posteriormente entregar a los responsables en la toma de decisiones información expresada de una forma que ellos pueden asimilar fácilmente. Esto se conoce como análisis a la velocidad del pensamiento, es decir, la capacidad de obtener una respuesta a una pregunta tan rápido como la pregunta es formulada. Esto permite hacer un gran salto en la calidad del análisis que puede ser ejecutado, trayendo como consecuencia un mejor entendimiento del negocio.

A pesar de que la tecnología es una parte importante, frecuentemente el aspecto más complejo del *Business Intelligence* es ser capaz de identificar qué información es útil y relevante para la toma de una decisión. Las soluciones de BI a un nivel corporativo son las responsables de identificar y cuantificar las métricas más importantes de la empresa, conocidas generalmente como indicadores de gestión o indicadores estratégicos de gestión (*KPI – Key performance indicators*). Los *KPI* sirven de orientación a la hora de tomar decisiones que afecten tanto a determinadas unidades de negocio como a la compañía en su conjunto.

Además de los *KPI*, ahora la visión de BI se ha expandido más allá de la medición interna, que tradicionalmente ha caracterizado a la administración de informes, con los avances de los sistemas de comercio electrónico, *Business-to-Business* (B2B) y *Business-to-Consumers* (B2C), *Business Intelligence* constituye cada vez una herramienta para suministrar información solicitada por personas de fuera de las organizaciones y a menudo una fuente de ingresos.

### Utilizar un método razonable para la gestión empresarial

BI puede ser definido como un método para la gestión empresarial.

Tanto las personas como las organizaciones se interesan en el BI, porque creen que el uso de un enfoque racional y basado en hechos a la hora de tomar decisiones resulta positivo en la medida que sea posible.

El interés de adoptar el *Business Intelligence* tiene las siguientes características:

- Buscar hechos (datos) que se puedan medir cuantitativamente acerca del negocio (organización).
- Usar métodos organizados y metodologías para analizar los hechos.
- Inventar o compartir modelos que expliquen las relaciones de causas y efecto entre las decisiones operativas y los efectos que éstas tienen en alcanzar los objetivos de negocio.
- Experimentar con métodos alternos y supervisar con retroalimentación los resultados
- Comprender que las personas no siempre son seres racionales.
- Gestión de la empresa (decisiones e iniciativas) basadas en todas estas características.

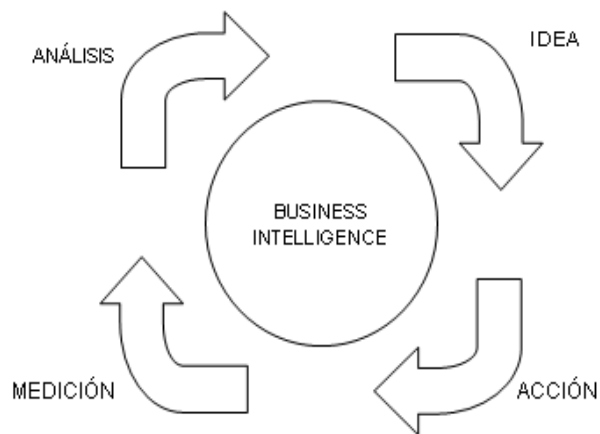
### **Racionalidad y ciencia.-**

La forma en que *Business Intelligence* aplica la racionalidad en la gestión de la empresa es una reminiscencia de la forma en que la ciencia utiliza la racionalidad para estudiar el universo. Muchos aspectos de la ciencia, tales como recopilar datos, dar forma y evaluar teorías y experimentar, tienen paralelos en la mayoría de las aplicaciones de *Business Intelligence* rigurosas. Mientras la ciencia busca estudiar un fenómeno aislado en escenarios experimentales meticulosamente controlados, *Business Intelligence* negocia con el comportamiento de los clientes, proveedores, competidores, empleados y otras entidades en la áspera e informal conducta de negocio del día a día. La ciencia pura busca el entendimiento por su propio bien sin una fecha límite. *Business Intelligence* busca el entendimiento con la finalidad de tomar decisiones para cumplir los objetivos organizacionales.

### El Ciclo de *Business Intelligence*

Es importante entender que *Business Intelligence* es más que una actitud empresarial o solo una tecnología, de hecho es un marco de referencia para la gestión del rendimiento empresarial, un ciclo continuo por el cual las compañías definen sus objetivos, analizan sus progresos, adquieren conocimiento, toman decisiones, miden sus éxitos y comienzan el ciclo nuevamente.

Normalmente se analizan datos provenientes de muchas fuentes. El análisis conduce a ideas, sugieren maneras de mejorar el negocio cuando se actúa sobre ellas; estas ideas pueden ser medidas para ver si funcionan. Estas mediciones también proveen más datos para el análisis, y el ciclo comienza de nuevo, a esta progresión (del análisis a la idea a la acción a la medición) la llamamos ciclo de *Business Intelligence*, como puede verse en la Figura 2.2.



**Figura 2.2 El ciclo de Business Intelligence.**

#### a) Análisis

El primer paso del ciclo de *Business Intelligence* es hacer y responder muchas preguntas rápidamente, convencionales y no convencionales. Esto es lo que el análisis real significa, la autonomía de tener curiosidad y hacer cientos de preguntas hasta que consigamos una respuesta brillante.

#### b) Ideas

La idea es el producto del análisis amplio, sin restricciones nacido de preguntas que solo el ser humano se puede hacer el descubrimiento de patrones que solamente los humanos pueden reconocer como útiles.

Una idea por sí sola no siempre es aceptada nada más porque esta sea brillante o incluso correcta. ¿Cuál es el significado de esto para *Business Intelligence*? Si una persona tiene una idea importante, esta generalmente tiene que ser compartida por otros para que sea de utilidad.

#### c) Acción

La acción se conecta con el ciclo de BI a través del proceso de toma de decisiones. La acción es lo que sigue de una mejor y más rápida toma de decisiones proporcionadas por el BI. Las decisiones bien fundamentadas (decisiones apoyadas en un buen análisis y sus conclusiones) proporcionan una motivación extra al que emprende la acción. Al contrario de tomar decisiones por el método de ensayo y error, que a menudo resultan en planes de acción limitados y en una menor financiación para los proyectos, la acción respaldada por una análisis sólido y por el *Business Intelligence* se caracteriza por unos objetivos más claros y en un mayor apoyo por parte de la organización para su implementación.

#### d) Medición

A través de una mejor recopilación de la información y una mayor frecuencia en la generación de informes que *Business Intelligence* proporciona, resulta posible medir los resultados de la compañía en comparación con los estándares cuantitativos, con la finalidad de desembocar en un nuevo ciclo de análisis, ideas y acciones correctivas.

El término empleado por *Business Intelligence* para designar las métricas importantes se denomina indicadores de gestión (*KPI – Key Performance Indicators*). Los sistemas de BI están específicamente diseñados para asimilar grandes cantidades de datos complejos con el fin de asignar, agregar, y en definitiva, jugar con la información. El resultado es la obtención sistemática de informes con la métrica, ratios e indicadores del negocio (esto es, los auténticos KPI) que los gerentes necesitan identificar, analizar y utilizar para tomar decisiones de forma frecuente. En un sistema de BI realmente completo, todas las áreas funcionales del

negocio debes estar incluidas en el ámbito de los KPI. El objetivo es que los gerentes gestionen aquello que es posible gestionar, y estos suelen ser los KPI. En la siguiente tabla se puede ver unos ejemplos de KPI.

<b>Ejemplos de KPI</b>		
Se muestran como ejemplo algunos de los KPI más usados		
<b>Área Funcional</b>	<b>KPI más comunes</b>	
<b>Operaciones</b>	Capacidad de utilización Unidades de SKU producidas % de unidades defectuosas Tiempo de espera Rendimiento Costo de inventario Rendimiento del proveedor a planificar	Rotación de inventario Desviación Número de unidades rechazadas % de ordenes planificadas por solicitud Devolución Precisión en inventario
<b>Ventas/Marketing al cliente</b>	Unidades vendidas Precio de venta promedio Importe por cliente Ventas por vendedor Número de clientes encuestados	Importe de ventas Número de clientes Número de productos por cliente Ítems por orden Ventas por empleado
<b>Finanzas</b>	Variación /Presupuesto % de variación % de beneficio Ventas diarias destacadas Ratio veloz Ratio de devolución	Variación/Proyección % de margen % de ventas Ratio actual Rotación de activo

**Tabla 2.2 Ejemplos de KPI.**

### **Como hacer posible el *Business Intelligence*.**

Actualmente, hasta las empresas más pequeñas tienen aplicaciones *Business Intelligence*, se pueden desarrollar complejos análisis en comparación con las grandes organizaciones generaciones atrás. Las organizaciones hoy pueden crear amplios sistemas corporativos de *Business Intelligence* que calculan y supervisan métricas sobre cada variable que sea importante para la gestión de la compañía.

¿Cómo es esto posible? La respuesta es la tecnología, que es el factor clave que más contribuye a que las aplicaciones *Business Intelligence* se hagan realidad. Las tendencias que posibilitan el desarrollo de aplicaciones BI son:

#### Potencia de procesamiento

Actualmente, las computadoras entregan más potencia en comparación a los supercomputadores de una década atrás. La potencia de cómputo de un microprocesador se dobla fielmente cada 18 meses (Ley de Moore), mientras los precios se han mantenido constantes. El impacto sobre *Business Intelligence* tiene 2 caras:

- Las técnicas de análisis de BI que necesitan millones de cálculos pueden ahora hacerse rápidamente en servidores de bajo costo, para posteriormente distribuirse a los usuarios finales por un costo bajo.
- Una potencia de procesamiento grande ha revolucionado la tecnología de visualización disponible, impresionantes gráficos de los datos, mostrados en monitores muy avanzados, imágenes que se pueden proyectar sobre la pared en presentaciones a una audiencia, impresión de estas vistas y con animación en tiempo real.

#### Almacenamiento

La capacidad de almacenamiento de los datos ha tenido un crecimiento exponencial, en comparación con la década pasada, mientras los precios se han mantenido constantes incluso caído. La cantidad de transacciones que es capturada, procesada y disponible para ser analizadas en aplicaciones *Business Intelligence* no tiene precedentes. Un analista de negocios puede cargar una copia entera de la base de datos del departamento de ventas en un computador portátil. Tan grande como la

aplicación de *Business Intelligence* pueda cargar y triturar con rápido procesado, el costo de almacenar grandes cantidades de datos es insignificante comparado con los beneficios del análisis.

### Tecnologías de redes

El crecimiento de Internet, la adopción extendida de exploradores web y el desarrollo de interfaces amigables han permitido que las organizaciones puedan conectarse virtualmente a sus empleados con las bases de datos de las aplicaciones *Business Intelligence*.

*Business Intelligence* ya no es más del dominio de un selecto grupo de usuarios o analistas en un departamento quienes trabajan con los datos en sus hojas de cálculo. Las tecnologías de redes pueden estar al alcance de cada gerente y empleado, quienes necesitan conocer que está sucediendo en su parte del negocio. Como Bernard Liautaud<sup>3</sup> explica, estamos entrando en una era de información democrática.

### Estándares

Han quedado atrás aquellos días cuando el hardware y software de un computador era comprado a un mismo fabricante. Los usuarios inexpertos pueden comprar software en numerosas compañías, instalar este software en sus computadores y tener un alto grado de confianza de que todo el sistema funcionará. Los estándares permiten esta interoperabilidad, la interoperabilidad de *Business Intelligence* está en crecimiento acelerado debido a los estándares de acceso a los datos relacionales y multidimensionales que apoyan a los sistemas *Business Intelligence*.

### Software *Business Intelligence*

La industria del software de *Business Intelligence* existe poco más de 10 años. Actualmente existen grandes compañías que ofrecen paquetes completos de productos para BI; estos paquetes incluyen: herramientas para la extracción, transformación y carga de datos (*ETL – extracting, transforming, and loading data*), grandes sistemas para el almacenamiento y agregación de datos, interfaces para el usuario final, todo esto por una fracción del costo de los sistemas de soporte de decisiones que aparecieron en la década de los 90. El impacto que la utilización que

---

<sup>3</sup> Presidente y Fundador de Business Objects, compañía de software líder mundial en business intelligence (BI)



estas herramientas puede tener en la eficiencia y eficacia de una organización es enorme.

### **La gente**

Entender el papel de la gente en el ciclo Business Intelligence permite a las organizaciones sistemáticamente crear ideas y convertir estas ideas en decisiones. Una manera para que las organizaciones puedan mejorar sus decisiones es tener la gente correcta tomando las decisiones. Lo que significa que un gerente de campo cerrara más cuentas debido a su experiencia, más que un analista con mucha información.

En los últimos años “*Business Intelligence* para las masas” ha sido una importante tendencia, y muchas organizaciones han dado grandes pasos para proporcionar complejas herramientas analíticas para grandes cantidades de usuarios.

Los sistemas de información pueden informar de lo que está pasando en la empresa, pero adquirir conocimiento de estos acontecimientos requiere la motivación de la gente y las organizaciones. Y para que esto suceda definitivamente se necesita a la gente.

### **La Cultura**

Una responsabilidad clave de los ejecutivos es configurar y administrar la cultura corporativa. La extensión para la cual la actitud *Business Intelligence* florece en una organización depende en gran parte de la cultura organizacional.

Adoptar interés hacia *Business Intelligence* no es simplemente escribir cualquier cosa en plan de negocios, hablar en reuniones de la compañía, para luego esperar a que algo suceda. Una actitud Business Intelligence tiene que ser respaldada por decisiones. Los ejecutivos no pueden esperar las mejores y más rápidas decisiones al menos que ellos inviertan en la tecnología y la gente.

### **“Pensar en grande y comenzar en pequeña escala”**

Significa utilizar áreas de oportunidad BI para guiar el camino hacia la creación de varios Data Marts, es la visión macro de las necesidades de información de diversas áreas funcionales y unidades de negocio de la organización balanceada con los pasos que ayudaran a materializarla.

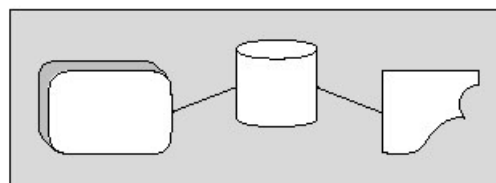
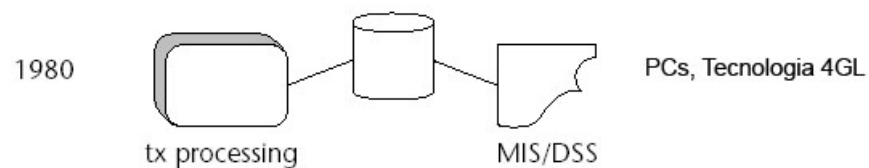
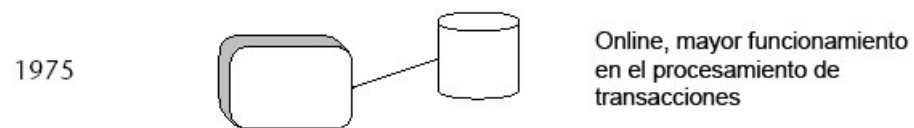
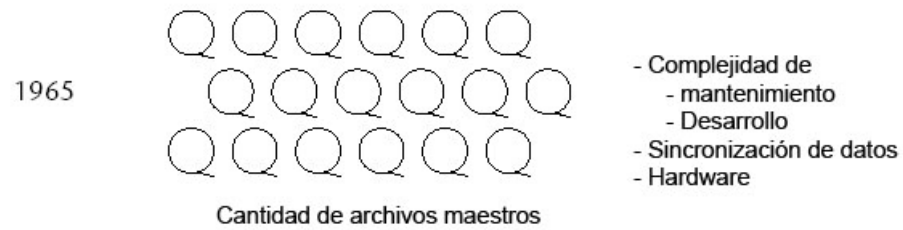
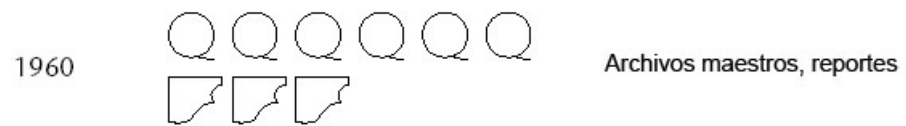
El objetivo global se traduciría en construir un Data Mart comenzando con el único que tenga la mayor prioridad y el nivel de esfuerzo menor para luego proceder con el siguiente Data Mart basado en su clasificación de importancia / dificultad.

### **2.3.2 Evolución de los Sistemas de soporte de decisión**

Como se vio anteriormente, el objetivo primario del *Business Intelligence* es ayudar a la gente a tomar decisiones, facultar a las organizaciones a tomar las mejores decisiones rápidamente.

En la figura 2.3 se muestra la evolución de los sistemas de soporte a la toma de decisiones desde los primeros días de la computación (1960) hasta los 80's.

A mediados de los 80's hicieron su aparición los sistemas OLTP, programas para la extracción de data y procesamiento de transacciones a través de una red de computadoras. El programa era el más simple de todos los programas, y trabajaba sobre archivos o BD.

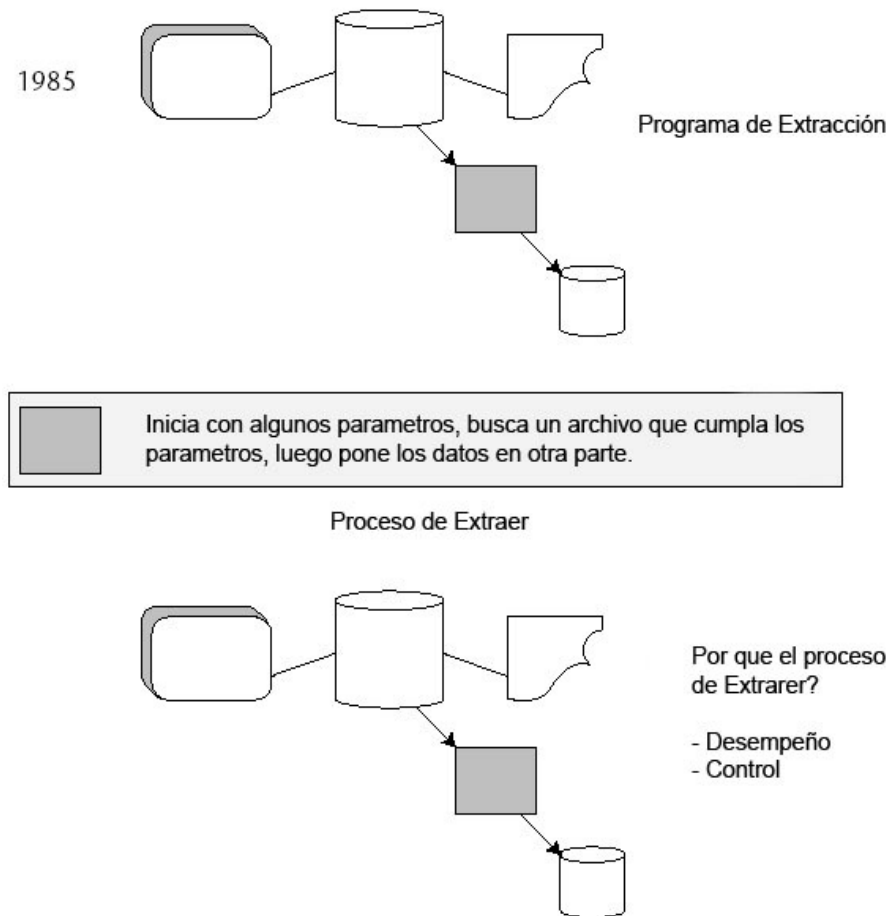


El paradigma de una sola base de datos sirviendo a todos los objetivos

**Figura 2.3 Las primeras etapas evolutivas del manejo de información.**

Como se muestra en la Figura 2.4, se comienza a formar una red.

Primero, había extracción de información; luego extracción de extracción; luego extracción de la extracción de la extracción; etcétera, etcétera. No era insólito para una empresa grande realizar no menos de 45,000 procesamientos por día.



**Figura 2.4. La naturaleza del proceso de Extraer.**

Este modelo de tratamiento de extracción de control a través de la organización se hizo tan banal que le dieron su propio nombre - " la arquitectura que se desarrolla naturalmente " - que ocurre cuando una organización maneja el proceso entero de hardware y la arquitectura de software con una actitud de política de no intervención.

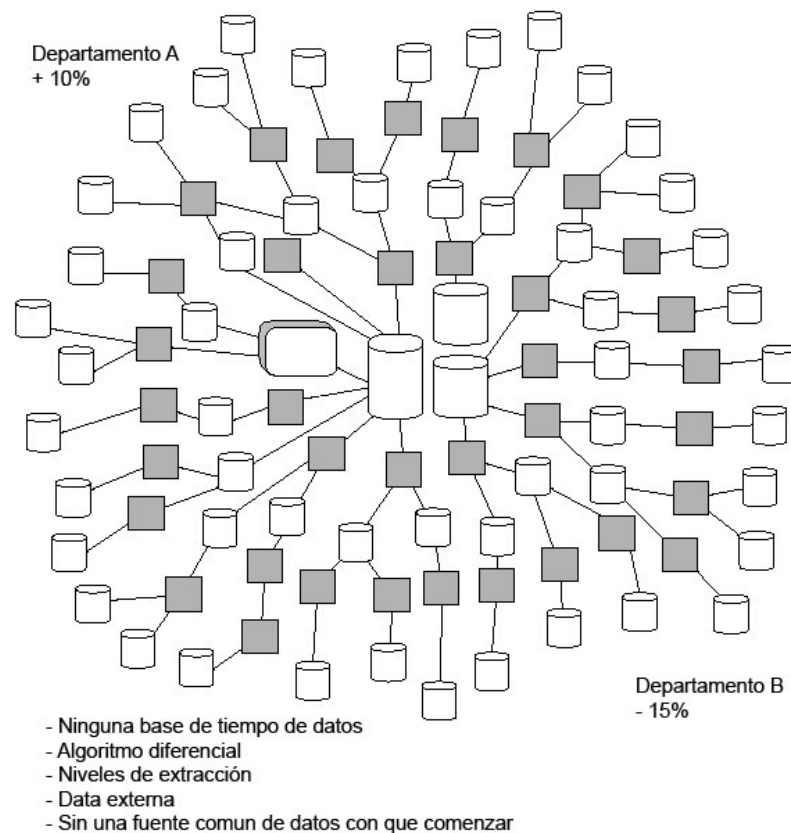
Más grande y más maduro la organización, pero los problemas de la arquitectura que se desarrolla naturalmente comenzaron a presentarse.

La arquitectura que se desarrolla naturalmente presenta muchos desafíos, como:

- Credibilidad de Datos.
- Productividad.
- Inhabilitado de transformar datos en información.

### ***Carencia de credibilidad de los datos***

La carencia de credibilidad de datos fue ilustrada en la Figura 2.5. Dos departamentos entregan un informe a la dirección, un departamento informa que la actividad bajo en 15 %, el otro dice que la actividad aumento en 10%. No sólo están los dos departamentos fuera de sincronización uno del otro, ellos están desconectados por márgenes muy grandes. Además, tratando de reconciliar la información diferente de los departamentos diferentes es difícil. A no ser que la documentación haya sido hecha muy cuidadosamente, la reconciliación es para todos los objetivos prácticos, imposibles.



**Figura 2.5 Falta de credibilidad en la arquitectura de crecimiento natural.**

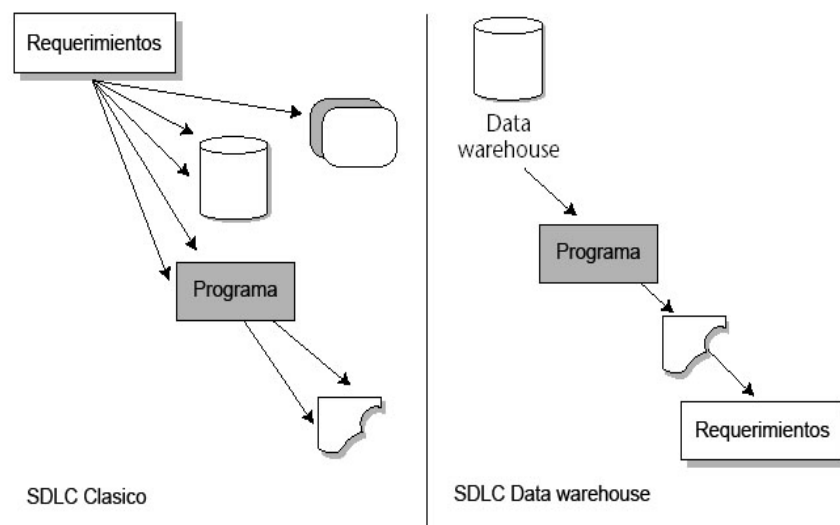
Cuando la dirección recibe los informes contradictorios, es forzado a tomar decisiones basadas en la política y personalidades porque ninguna fuente es más o menos creíble. Éste es un ejemplo de la crisis de credibilidad de datos en la arquitectura que se desarrolla naturalmente.

### 2.3.3 Data Warehouse

#### El desarrollo del ciclo de vida

Los datos operacionales por lo general son orientados por uso y, como consecuencia no son integrados, mientras que los datos del Data Warehouse deben ser integrados.

Otras diferencias principales también existen entre el nivel operacional de datos y tratamiento. Los ciclos de vida de desarrollo subyacentes de estos sistemas pueden ser de profundo interés, como se muestra en la figura 2.6



**Figura 2.6 El Ciclo de Vida de Desarrollo de Sistemas para el ambiente Data Warehouse es casi lo opuesto a el clásico SDLC.**

#### 2.3.4 MIGRACIÓN DE DATOS: EXTRACCIÓN, TRANSFORMACIÓN Y CARGA

La migración de los datos desde las fuentes operacionales al Data Warehouse requiere la necesidad de procesos para extraer, transformar y cargar los datos, actividad que se conoce como ETL.

La mayoría de los datos de origen son los datos operacionales actuales, aunque parte de ellos pueden ser datos históricos archivados.

Si los requerimientos de datos incluyen algunos años de historia es necesario desarrollar tres conjuntos de programas ETL: una Carga Inicial, una Carga Histórica, y una Carga Incremental.

- **Carga Inicial**

La carga inicial se asemeja mucho al proceso de conversión entre sistemas que se da en las organizaciones cuando pasan, por ejemplo, de sus viejos sistemas operacionales a un producto ERP.

- **Carga Histórica**

Este proceso debe verse como una extensión de la carga inicial, pero la conversión aquí es un poco diferente porque los datos históricos son datos estáticos.

A diferencia de los datos operacionales, los datos estáticos ya se archivaron en dispositivos de almacenamiento offline. Es común que con el transcurso del tiempo se eliminen elementos de datos que ya no sirven, se agreguen nuevos, se modifiquen los tipos de ciertos datos o los formatos de los registros, lo que implica que los datos históricos no necesariamente se puedan sincronizar con los datos operacionales. Por lo tanto los programas de conversión escritos para la carga inicial quizá no sean aplicables a la carga de datos históricos sin algunos cambios previos.

- **Carga Incremental**

Una vez que el Data Warehouse está cargado con datos iniciales e históricos, hay que desarrollar otro proceso para la carga incremental, que se ejecutará mensual, semanal o diariamente. Existen dos formas de diseñar la carga incremental:

- *Extraer todos los registros*: Se extraen todos los registros operacionales, independientemente de los valores que hayan cambiado desde la última carga realizada. En general esta opción no es viable debido al volumen de los datos, por eso la mayoría opta por la siguiente opción.
- *Extraer Deltas solamente*: Solo se extraen registros nuevos o registros que contengan valores que cambiaron desde la última carga realizada.

Diseñar programas ETL para extracciones delta es más fácil cuando las fuentes consisten en bases de datos relacionales y contamos con una columna “timestamp” para determinar los deltas.

Expliquemos ahora, lo que debe contemplar este proceso:

#### **A. Extraer los Datos.**

Que consiste en determinar técnicas, para combinar eficiencia en el uso de la data de origen, así como detectar redundancias y datos y algún otro ruido. Además, hay que distinguir un dato que puede estar duplicado en distintas tablas.

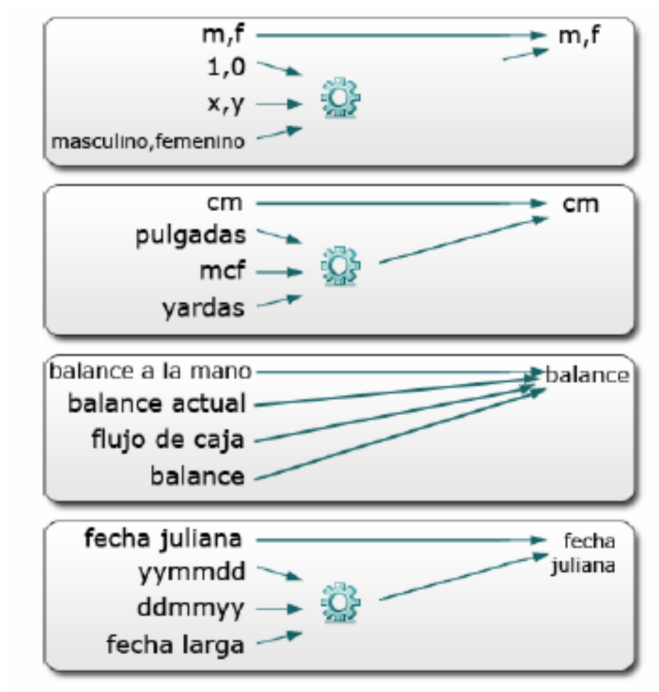
#### **B. Transformar Datos.**

Este proceso es el más crítico, debido a que debe controlar algunos factores:

Claves primarias inconsistentes, valores inconsistentes, datos con diferentes formatos, valores erróneos, sinónimos y homónimos, Lógica embebida, Integración y Derivación, etc.

En la figura vemos algunos ejemplos de transformación de datos: El primero referente a sexo, el segundo referente a unidades de medida, el tercero se refiere a estandarizar nombres, y por último, estandarizar formatos de fecha. En la figura 2.7 podemos ver un ejemplo de transformación.





**Figura 2.7. Ejemplos de Transformación.**

### **C. Cargar Datos.**

Este paso es el más simple, y sería el que completaría el proceso ETL. Aquí se tendría que tener cuidado, básicamente con los índices, y a la integridad referencial.

#### **2.3.5 Data Mart**

Un Data Mart es una solución que compartiendo tecnología con el Data Warehouse (pero con contenidos específicos, volumen de datos más limitado y un alcance histórico menor), permita dar soporte a una empresa pequeña, o un departamento o área de negocio de una empresa grande.

El Datamart cubre las necesidades de informes bien porque no es conveniente efectuar consultas sobre el sistema transaccional, bien porque hay que integrar datos de varios sistemas transaccionales.

Es un pequeños Data Warehouse, para un determinado número de usuarios, para un área funcional, específica de la compañía. También podemos definir que un Data Mart es un subconjunto de una bodega de datos para un propósito específico.

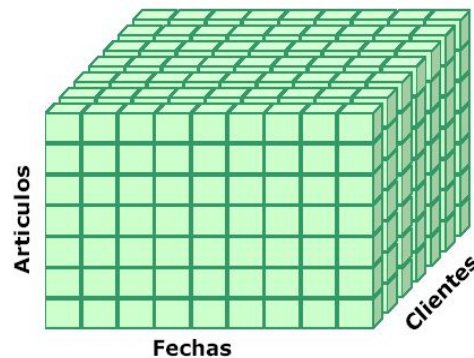
Su función es apoyar a otros sistemas para la toma de decisiones.

### 2.3.6 Bases De Datos Médica

Viene a ser la información de la historia clínica del paciente almacenados en soportes magnéticos o de cualquier otro material y susceptibles de ser tratados, recuperados y transmitidos de forma total o parcial mediante procedimientos y medios informáticos, que con la aplicación de técnicas derivadas de la utilización de la informática pretenden ser utilizados con una finalidad de registro local.

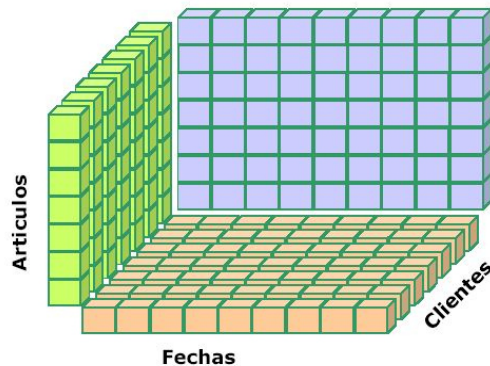
### 2.3.7 Cubo de Información

Es una estructura de datos multidimensional que representa la intersección de una combinación única de dimensiones. Para cada intersección hay una celda que contiene un valor. Un ejemplo de cubo de inf. se muestra en la figura 2.8.



**Figura 2.8: Cubo de datos.**

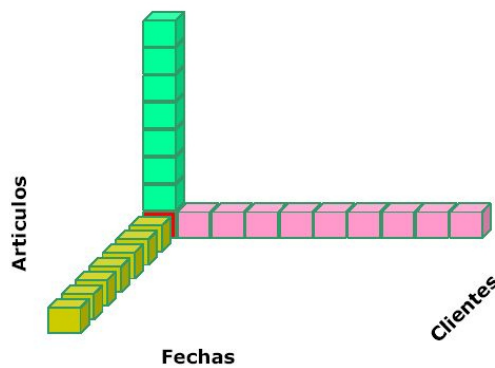
Una estructura de datos como la anterior admite una representación espacial en tres dimensiones, como se muestra en la figura 2.9. Cada cubo elemental representa una ocurrencia (fila) en la tabla de hechos.



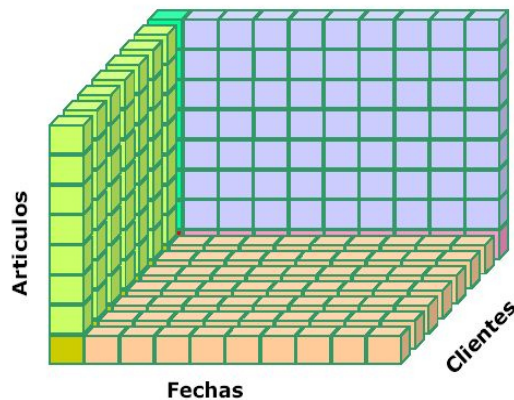
**Figura 2.9: Ocurrencias en la tabla de hechos.**

Las medidas (como el beneficio) tiene la propiedad de ser aditivas. Es decir, tiene sentido la suma según todas las dimensiones (beneficios en una fecha, o con un artículo o con relación a un cliente).

Además de almacenar los valores elementales de las medidas, se pueden también guardar los acumulados según las dimensiones. Un ejemplo se muestra en la figura 2.10. Y un ejemplo de acumulado de todas las dimensiones se muestra en la figura 2.11.



**Figura 2.10: Acumulados según dos y las tres dimensiones.**



**Figura 2.11: Acumulados según todas las dimensiones.**

Tamaño máximo del cubo:

$$(n_1+1) (n_2+1) \dots (n_d+1)$$

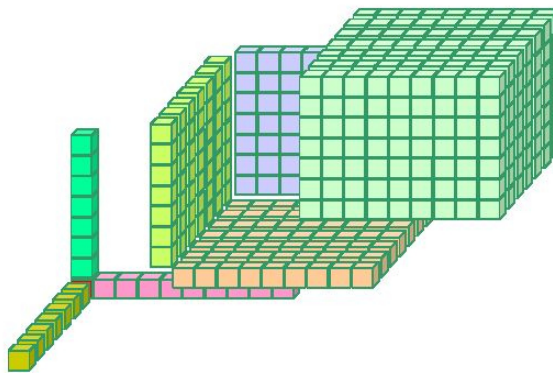
Donde:

$N_i$ : es el número de filas de la dimensión  $i$ .

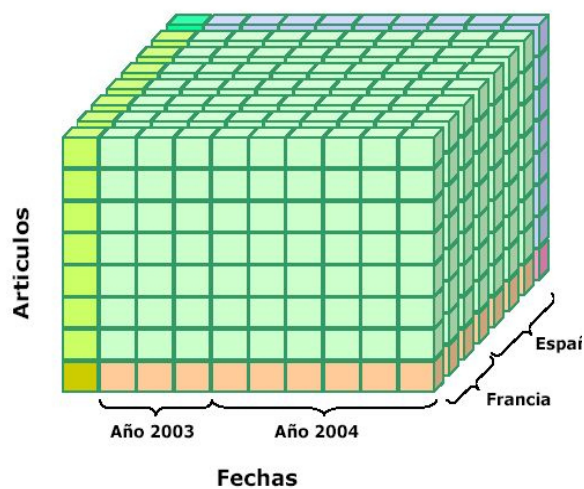
$d$ : es el número de dimensiones.

Muchos de los cubos no tiene medida (no ocupan espacio).

La unión del cubo y sus acumulados (Figura 2.12), dan como resultado un cubo de medidas con todos sus acumulados como se muestra en la Figura 2.13.



**Figura 2.12: Unión del cubo y sus acumulados.**

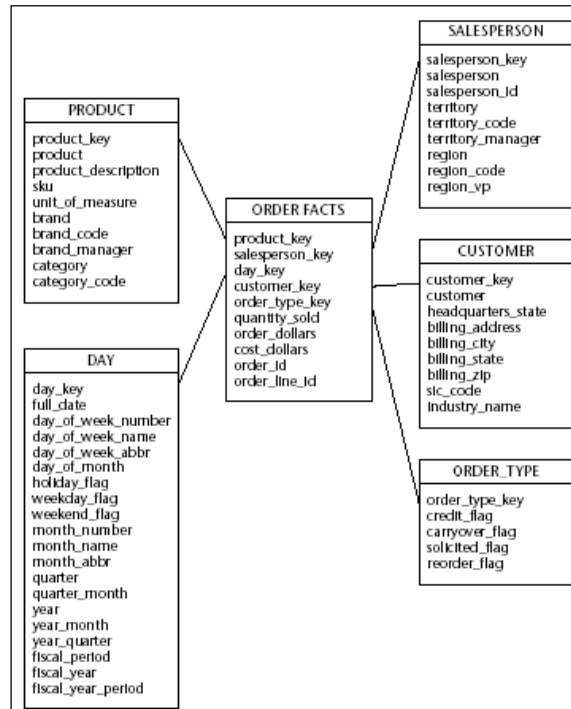


**Figura 2.13: El cubo de medidas con todos sus acumulados.**

### 2.3.8 Modelo Dimensional <sup>4</sup>

En un modelo dimensional, cada grupo de dimensiones es colocado en una tabla de dimensión; los hechos son colocados en una tabla de hecho. El resultado es un esquema o modelo de estrella, porque esto se parece a una estrella cuando se hace el diagrama con la tabla de hecho en el centro.

Un esquema de estrella para el proceso de órdenes es mostrado en la figura 2.14



**Figura 2.14: Modelo estrella para procesar órdenes.**

<sup>4</sup> Mastering Data Warehouse Aggregates: Solutions for Star Schema Performance, Christopher Adamson .

Las tablas de dimensión en un modelo estrella son amplias. Ellos contienen un número grande de atributos que proporcionan datos importantes para apoyar una amplia variedad de informes y análisis.

En la figura 2.15 vemos las diferencias entre el modelo dimensional y le modelo relacional.

Modelo		
	Relacional	Dimensional
<b>Objetivos</b>	Actualización de los datos	Consultas estratégicas
<b>Datos</b>	Actualizados dinámicamente	Históricos estáticos
<b>Usuario</b>	Sólo consultas elementales	No puede actualizar datos
<b>Consistencia</b>	Se persigue	Se da por supuesta
<b>Redundancias</b>	Se impiden	Se permiten
<b>Consultas estr.</b>	Costosas y presentación difícil	Sencillas y presentación fácil
<b>Diseño</b>	Alejado del usuario final	Próximo al usuario final

**Figura 2.15 Diferencias entre el modelo dimensional y el modelo relacional.**

### 2.3.9 DEFINICIÓN DE TÉRMINOS BÁSICOS

ETL.- Son las siglas en inglés de Extraer, Transformar y Cargar (*Extract, Transform and Load*). Proceso que permite a las organizaciones mover datos desde múltiples fuentes, reformatearlos y limpiarlos, y cargarlos en otra base de datos, Datamart, o Data warehouse para analizar, o en otro sistema operacional para apoyar un proceso de negocio.

OLAP.- Es el acrónimo en inglés de procesamiento analítico en línea (*On-Line Analytical Processing*). Es una solución utilizada en el campo de la Inteligencia de Negocios (*Business Intelligence*), la cual consiste en consultas a estructuras multidimensionales (o Cubos OLAP) que contienen datos resumidos de grandes Bases de Datos o Sistemas Transaccionales (OLTP). Se usa en informes de negocios de ventas, marketing, informes de dirección, minería de datos y áreas similares.

La razón de usar OLAP para las consultas es la velocidad de respuesta. Una base de datos relacional almacena entidades en tablas discretas si han sido normalizadas. Esta estructura es buena en un sistema OLTP pero para las complejas consultas multitabla es relativamente lenta. Un modelo mejor para búsquedas, aunque peor desde el punto de vista operativo, es una base de datos multidimensional. La principal característica que potencia a OLAP, es que es lo más rápido a la hora de hacer *selects*, en contraposición con OLTP que es la mejor opción para *INSERTS, UPDATES Y DELETES*.

P.O.O. .- Programación Orientada a Objetos, es un paradigma de programación que define los programas en términos de clases de objetos, objetos que son entidades que combinan estado (es decir, datos), comportamiento (ésto es, procedimientos o métodos) e identidad (propiedad del objeto que lo diferencia del resto). La programación orientada a objetos expresa un programa como un conjunto de estos.

Análisis Multidimensional (Multidimensional análisis).- Es una forma de análisis de datos en un estilo arriba – abajo (top – down) examinando medidas simultáneamente agrupadas en múltiples dimensiones.

Base de Datos (Database).- Es un conjunto de datos relacionados y organizados de una forma útil para fácil recuperación. Existen diferentes tipos de bases de datos dependiendo del tipo de datos que están almacenados y como usarlos.

Bases de datos OLAP (Online Analytical Processing (OLAP)).- Es una base de datos que permite análisis multidimensional. El análisis multidimensional está compuesto por herramientas de interfaces y estructuras de bases de datos que permiten acceso instantáneo y una fácil manipulación por parte del usuario. OLAP obtiene su nombre por un buen contraste con OLTP, un término que ya estaba ampliamente usado cuando el termino OLAP fue creado. Existen diferencias fundamentales entre procesamiento de transacción y procesamiento analítico. Los sistemas OLAP contienen análisis multidimensional a la velocidad del pensamiento, OLAP típicamente es la continuación del paradigma “cliente / servidor” donde es alcanzado por muchos usuarios que usan herramientas de clientes multidimensionales para analizar la información.

Base de datos Operacional (Operational database).- Una base de datos que contiene las operaciones diarias de una organización. Las bases de datos operacionales albergan los sistemas que las organizaciones utilizan a diario para ejecutar sus procesos de negocio. La mayoría de las bases de datos operacionales son sistemas OLTP y almacenan la información en base de datos relacionales.

Bases de datos MOLAP (Multidimensional Online Analytical Processing).- Es una base de datos OLAP en la cual es colocada en estructuras especiales que son almacenadas en un servidor central.

Bases de datos ROLAP (Relational Online Analytical Processing).- Es un modo de almacenamiento OLAP donde la data es almacenada en bases de datos relacionales.

BI.- Acrónimo de *Business Intelligence*

Brecha en el Análisis (Análisis Gap).- Es la brecha que existe entre la información necesaria para la toma de decisiones en una empresa y los datos que son recopilados por los sistemas de información empresarial.



Business-to-Business (B2B).- En forma breve, se define como el intercambio de bienes, servicios, productos o información entre empresas.

Business-to-Consumers (B2C).- Es un marco de referencia para la administración del desempeño; un ciclo en marcha por medio del cual las compañías definen sus objetivos, analizan sus procesos, ganan en conocimiento, actúan , miden sus éxitos y vuelven a empezar.

Dimensión.- Es una visita de datos categóricamente consistente. Todos los miembros de una dimensión pertenecen a un mismo grupo.

Interoperabilidad.- Es la propiedad o capacidad de un producto para trabajar juntos e interactuar con otros productos.

Modelo Mental.- Un conjunto de todo lo que creemos que conocemos acerca de cómo son las cosas (nuestro negocio). Esta denominación de nuestro entendimiento se aplica no solo a las personas sino también a organizaciones.

Slice and Dice.- Son dos métodos complementarios para interactuar con los datos. *Slicing* “rebanar” significa aislar un miembro específico de una dimensión para hacer análisis. *Dicing* significa dividir o romper un conjunto de datos en pequeñas piezas para examinar como las medidas intersectan múltiples dimensiones.

Metástasis Ganglionar: Las metástasis ganglionares se producen habitualmente en el sentido de la corriente linfática, por lo cual aparecen en general ordenadamente, primero en los ganglios que drenan el territorio del tumor primario y así sucesivamente, alejándose.

Adenocarcinoma: Cáncer que se inicia en la células que revisten ciertos órganos internos y que tienen propiedades glandulares (secretorias).

## **CAPÍTULO 3: ESTADO DEL ARTE**

Las organizaciones que cuentan con sistemas de apoyo a la toma de decisiones agilizan las operaciones que sostienen a una empresa. A pesar de la riqueza de estos datos, no se puede recurrir a ellos con facilidad cuando el médico oncólogo necesita tomar decisiones para contrarrestar los problemas que se presentan después de la cirugía, como por ejemplo en el caso del INEN:

- ¿Cuál es la cantidad de pacientes, tuvieron una mortalidad hospitalaria o muerte a 30 días?
- ¿Cuál es la morbilidad post-operatoria de un determinado diagnóstico de cáncer de páncreas en los últimos 5 años?
- ¿Qué doctores han tenido una mayor cantidad de complicaciones post-operatorias han surgido en los últimos 5 años?

### **3.1. DISTRIBUCIÓN A NIVEL NACIONAL**

Las enfermedades del aparato digestivo ocupan en el Perú el segundo lugar en mortalidad y los tumores malignos del aparato digestivo el tercer lugar, por lo que se planteó conocer las tasas de mortalidad de cada una de las enfermedades digestivas y hepatobiliares, así como establecer su frecuencia y distribución por áreas geográficas en el país.

Para establecer un adecuado enfoque y orientación en la salud individual y colectiva de un país, es necesario conocer el perfil epidemiológico de las enfermedades que la afectan. Este es el caso de las enfermedades digestivas, en las que es necesario investigar y conocer la prevalencia, incidencia, factores de riesgo y tendencias a través del tiempo.

Al respecto en el Perú, existen estudios sobre la distribución de enfermedades digestivas y hepato-biliares realizados por diferentes autores en determinadas ciudades o departamentos, muchos de ellos basados en los registros de hospitales, los cuales pueden servir como referencia para orientar otros estudios más exhaustivos. La información basada en la mortalidad por ciudades o regiones, son las de mayor utilidad para el análisis, y más aún si son casos determinados con residencia habitual en la ciudad correspondiente, además de su distribución por grupos de edad y género.

Existe información del nivel nacional en el Ministerio de Salud cuya selección y catalogación ha variado en los últimos años, por otro lado hay estudios de prevalencia de determinadas enfermedades digestivas que pueden correlacionarse con las anteriores. Es importante hacer notar que las primeras cinco causas de mortalidad y morbilidad, según diferentes series, representan alrededor de dos tercios de las defunciones totales [Farfan00].

Basados en los datos disponibles, en el presente estudio, la información fue procesada con los objetivos de: calcular las tasas de mortalidad de las diferentes enfermedades digestivas y hepato-biliares en el Perú, de los años 1995 al 2000, estableciendo las cinco primeras causas por cada tipo de enfermedad y su distribución en cada departamento, así como su distribución por género y grupos de edades, evaluar estos resultados asociados a otro tipo de estudios. De modo que esta información pueda orientar las necesidades de información adicional y se puedan diseñar pautas para la prevención y control de las enfermedades digestivas en el país.

### **3.2. TÉCNICAS Y MÉTODOS**

Se obtuvo información de los diagnósticos por certificados de defunciones, en la Oficina de Estadística e Informática de Instituto Nacional de Enfermedades Neoplásicas (INEN) de los años 1995 a 2000. Esta información fue revisada, para elegir determinadas enfermedades y luego ser catalogadas como tasas de mortalidad (TM), por poblaciones o departamentos, relacionándola con información poblacional de la misma fuente del MINSA y determinando las tasas por cada 100,000 habitantes del departamento por cada año.

Las tasas obtenidas se ordenaron en cifras generales por cada año, divididas en tres grupos: enfermedades digestivas no tumorales (patologías esofagogástricas, vesiculares,

pancreáticas y colon), enfermedades tumorales (esofágicas, gastrointestinales, páncreas, vesícula y vías biliares y de hígado y vías biliares intrahepáticas) y enfermedades hepáticas (cirrosis hepática, tumor hepático maligno primario y hepatitis virales agudas).

De las tasas generales, a nivel nacional, por cada enfermedad, se agruparon las poblaciones por grupos de edades y por género masculino y femenino.

Toda la población del presente trabajo elaborado se basa en pacientes atendidos en el Departamento de Abdomen, y a los cuales se les realizó pancreatoduodenectomía en el Departamento de Abdomen del Instituto Nacional de Enfermedades Neoplásicas (Enero 2002 – Enero 2007).

Debido a la insuficiente cantidad de datos digitalizada para la realización de este trabajo, se tuvo que extraer la información de las historias clínicas, basándose en una ficha de un trabajo de investigación, elaborándose también en el INEN. El formato de la ficha se muestra a continuación [Valencia08].

**FÍSTULA PANCREÁTICA EN FUNCIÓN AL CONTROL SISTEMÁTICO DE AMILASA EN  
PACIENTES SOMETIDOS A DUODENOPANCREATECTOMÍA EN EL INSTITUTO DE  
ENFERMEDADES NEOPLÁSICAS. 2002-2007**

No Registro  Historia  No

Edad  Sexo  Hombre  Mujer

**Estado preoperatorio:**

**ASA:** ( I ) ( II ) ( III ) ( IV ) ( V ) **RQ:** ( I ) ( II ) ( III ) ( IV )

**Zubrod** /(0) (1) (2) (3) (4) (5)

**Patologías asociadas:**

**Diabetes mellitas:** Si ( ) No ( ) **HTA:** Si ( ) No ( )

**Asma:** Si ( ) No ( ) **Coronariopatía:** Si ( ) No ( )

**Colangitis:** Si ( ) No ( ) **Cirugía previa:** Si ( ) No ( )

**Otros:** \_\_\_\_\_

**Laboratorio (pre cirugía):**

**Hgb:** \_\_\_\_\_ g/l **Alb:** \_\_\_\_\_ g/l

**Brrb:** **Tot** \_\_\_\_\_ umol/l **Fa:** \_\_\_\_\_ u/l  
**Dir** \_\_\_\_\_ umol/l  
**Indir** \_\_\_\_\_ umol/l

**Tgo:** \_\_\_\_\_ **Tgp:** \_\_\_\_\_

**Drenaje biliar pre QX:** Si ( ) No ( ) **PCRE-STENT** **CPT**

**Datos quirúrgicos:**

**Fecha:** \_\_\_\_ / \_\_\_\_ / \_\_\_\_

**Cirugía:** Pancreatoduodenectomía clásica ( )  
Pancreatoduodenectomía preservadora de píloro ( )

**Cirujano:**

**Higado de aspecto:** Normal Estratos. Cirrótico

**Tamaño tumoral:**

**Sospecha metástasis ganglionar** Si ( ) No ( )

**Infiltración Estructuras adyacentes:**

**Sangrado Operat.:**

**Tiempo Operatoria:**

**Textura páncreas:** Blando ( ) duro ( )

**Octeotride Postopera.:** Si ( ) No ( )

**Diam. wirsung:** \_\_\_\_\_ mm.

**Cultivo bilis:** Si No

Germen aislado	Sensibilidad

	PostQx	1	2	3	4	5	6	7	8	9	10	11	12
<b>Der</b>	<b>A</b>												
	<b>V</b>												
	<b>Se Sh</b>												
	<b>Si</b>												
	<b>S</b>												
	<b>B</b>												
<b>Izq</b>	<b>A</b>												
	<b>V</b>												
	<b>Se Sh</b>												
	<b>Si</b>												
	<b>S</b>												

	<b>B</b>												
<b>Muestra</b>	<b>Leu</b>												
	<b>Abast</b>												
	<b>Amil sér</b>												
	<b>Vol</b>												
<b>Yeyuno</b>	<b>Biliar</b>												
	<b>Intest.</b>												
	<b>Sang.</b>												
	<b>Otro</b>												

**Morbilidad postoperatoria:**

Fístula pancreática ( )  
 Retardo vaciamiento gástrico ( )  
 Colección intra abdominal ( )  
 Hemorragia gastrointestinal ( )  
 Hemorragia post operatoria intra abdominal ( )  
 Fístula biliar ( )  
 Ileo ( )  
 Neumonía ( )  
 Complicaciones de herida operatoria ( )  
 Infarto de miocardio ( )  
 Distress respiratorio ( )

**Resop:** Si ( ) No ( ) **Motivo:** \_\_\_\_\_

**UCI:** Si ( ) No ( )

**Fecha ingreso:**            /        /                      **Fecha alta:**            /        /

**Cultivos:**

<b>Fecha</b>	<b>Tipo Cultivo</b>	<b>Germen aislado</b>
--------------	---------------------	-----------------------


**Muerte** Si ( ) No ( )  
**30 días:** )

**Muerte** Si ( ) No ( )  
**Hosp.:** )

---

**Diagnóstico Patológico:**

Cáncer de páncreas	( )	EC	T	N	M	para
	NM					
Neoplasia ampular	( )					
Neoplasia de coledoco distal	( )					
Neoplasia de duodeno	( )					
Cistoadenoma mucinoso	( )					
Neoplasia mucinosa papilar intraductal	( )					
Cistoadenoma seroso	( )					
Cistoadenocarcinoma	( )					
Otros						

---

**Control y seguimiento:**

**Recurrencia:** Si ( ) No ( )

**Fecha de último control:**        /        /

**Estado:** Vivo ( )  
muerto ( )



Y la definición de las variables usadas en esta ficha, se muestra a continuación:

<b>Variable</b>	<b>Definición</b>	<b>Tipo variable</b>	<b>UM</b>	<b>Categorías</b>
Teléfono	Número telefónico consignado en HC			
Dirección	Domicilio consignado en HC			
Edad	Edad cronológica considerada en años cumplidos al momento de la cirugía	Continua		
Sexo	Sexo del paciente	Nominal		Hombre Mujer
Fecha_ingreso	Fecha en que se hospitaliza el paciente para tratamiento quirúrgico.			
Fecha_alta	Fecha en que es dado de alta el paciente luego de tratamiento quirúrgico, considerando los reingresos hasta el postoperatorio 30			
Drenaje	Procedimiento de drenaje de la vía biliar	Nominal		Si, no
PCRE-STENT	Realización previa a la cirugía de pancreatocolangiografía retrograda endoscópica hasta 2 meses antes o presencia del stent en el momento de la cirugía	Nominal		Si No
CPT	Colangiografía percutaneo transcoledociana	Nominal		Si No
ASA	Graduación del estado preoperatorio valores del I al V	Nominal		I II III IV V
RQ (Riesgo Quirúrgico)	Estado preoperatorio	Nominal		I II III IV
Fecha_cirugía	Fecha de acto quirúrgico			
Tiempo_operatorio	Tiempo operatorio	Continua	M	
Pérdida_sangre	Pérdida sanguínea durante el acto quirúrgico	Continua	cc.	
Textura_páncreas	Textura del páncreas	Nominal		Blando Duro
Octeotride	Si se utilizó octeotride como profilaxis en el postoperatorio.	Nominal		Si No
Diámetro_wirsung	Diámetro de wirsung (mm.)	Continua	mm	

A	Cantidad de amilasas valor en secreción drenada (drenes derecho e izquierdo y post operatorios 1, 3, 6, 9 hasta 12)	Continua	U/ml	
V	Volumen en cc. en 24 horas (drenes derecho e izquierdo y post operatorios 1, 3, 6, 9 hasta 12)	Continua	Cc	
Color	Color evaluado de manera subjetiva (drenes derecho e izquierdo y post operatorios 1, 3, 6, 9 y 12)	Nominal		Serohemático Turbio Purulento Sanguinolento Biliar
Uti	Permanencia del paciente en manejo en UTI. (Unidad de Tratamientos Intensivos)	Nominal		Si No
Fecha_ingreso	Ingreso unidad de cuidados intensivos			
Fecha_alta	Salida de unidad de cuidados intensivos			
Retardo_vaciamento	Definida según el consenso de ISGPF (ver anexo – Tabla 1)	Nominal		
Fístula_pancreática	Definida según el consenso de ISGPF (ver anexo – Tabla 2)	Nominal		
Absceso_intra	Cultivo positivo de colección intraabdominal obtenido por intervencionismo o de manera quirúrgica, con signos sistémicos o locales de infección.	Nominal		Si No
Hemorragia_gastro	Hemorragia gastrointestinal diagnosticada por endoscopia digestiva alta o baja, hematemesis, hematoquesia, melena caída de hb de 2 mg/dl. que requiere transfusión sanguínea o reintervención.	Nominal		Si No
Hemorragia_post	Definida según el consenso de ISGPF (ver anexo – Tabla 3)	Nominal		
Fistula_biliar	Secreción con bilirrubinas elevadas a nivel de drenajes de lecho quirúrgico o en colección intra abdominal confirmada radiográficamente que requiere cirugía, endoscopia o intervencionismo.	Nominal		Si No
Com_her_op	Signos de infección que requieren tratamiento antibiótico o evidencia de dehiscencia.	Nominal		Si No
Neumonía	Nuevo infiltrado en la rx tórax, con cultivo positivo de aspirado o esputo, con fiebre o leucocitosis que requiere antibióticos	Nominal		Si No

	endovenosos.			
Ileo	Ausencia de flatos, ruidos hidroaéreos o movimiento intestinal luego del PO 5 que requiere nutrición parenteral.	Nominal		Si No
Inf_miocardio	Paciente que cursa con incremento de CPK MB o troponina, y cambios en el EKG.	Nominal		Si No
Itu	Cultivo de orina positivo o piuria que requirió tratamiento antibiótico.	Nominal		Si No
Embo_pulm	Episodio de disnea, taquicardia o, hipotensión con incremento de dímero D cuantitativo y evidencia radiológica de tromboembolismo pulmonar.	Nominal		Si No
Comp_neurolo	Hipoxia cerebral, accidente cerebrovascular o hemorragia intracraneana con síntomas neurológicos.	Nominal		Si No
Distress_respirat	Incremento del PCO2 >60mmhg que requirió terapia farmacológica o intubación orotraqueal, por más de 24 hs en el postoperatorio.	Nominal		Si No
Resop	Reingreso a sala de operaciones para solucionar complicación postoperatoria.	Nominal		Si No
UCI	Manejo de la complicación en UCI desde el PO 1 no se considera los casos en que el paciente pase directamente a UCI.	Nominal		Si No
Fecha_ingreso	Ingreso unidad de cuidados intensivos.			
Fecha_alta	Salida de unidad de cuidados intensivos.			
Mortalidad_30_dias	Mortalidad debido a complicación quirúrgica durante la hospitalización inicial o luego de 30 días del alta.	Nominal		Si No
Diagnóstico	Diagnóstico patológico final.	Nominal		Cáncer de páncreas Neoplasia ampular Neoplasia de colédoco distal Neoplasia de duodeno Cistoadenoma mucinoso Neoplasia mucinosa papilar intraductal

				Cistoadenoma seroso Cistoadenocarcino ma Otros
Recurrencia	Recurrencia de enfermedad	Nominal		Si No
Fecha_control	Fecha de último control			
Estado	Estado del paciente en la fecha del último control	Nominal		Vivo Fallecido
Observación	Indicar situaciones especiales no consideradas en los ítems previos de relevancia para el investigador.			

Tabla 3.1: TABLA DE VARIABLES

Una vez obteniendo la información requerida desde las fichas: “Fístula Pancreática en función al control sistemático de amilasa en pacientes sometidos a duodenopancreatectomía en el Instituto Nacional de Enfermedades Neoplásicas 2002 - 2007”, se inicio la elaboración del Diagrama Funcional (Star Net) para la identificación de las jerarquías y dimensiones de nuestro data mart, como se muestra en la figura 3.1.

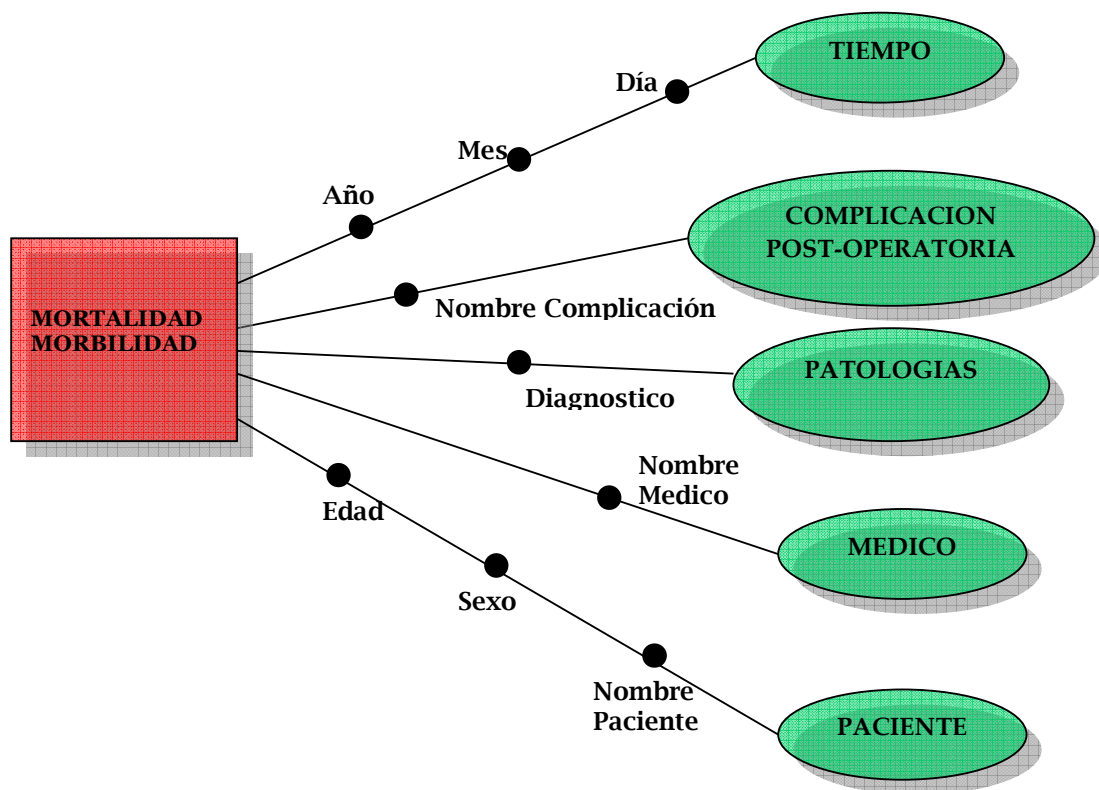


Figura 3.1: Diagrama Funcional (Star Net).

### 3.3. CASO DE ESTUDIO:

#### Ejemplos de elaboración de Data Mart.

#### CASO 1: DATAMART PARA UN CONSULTORIO

Para ver de manera general las fases de un Data Mart, en la figura 3.2 se presenta un ejemplo de Data Mart para un Consultorio, donde es el usuario quién genera la información (a través de archivos y sistemas transaccionales) que será usada como fuente del ETL (extraer, transformar y cargar la información de las fuentes de información). Se elaboró el modelo multidimensional y la tabla de hechos para luego representarlos en estructuras multidimensionales (cubos).

En la figura 3.2 se observa desde el proceso ETL hasta el uso de las herramientas OLAP y como se muestra los resultados que requiere el usuario en forma de tablas o reportes.

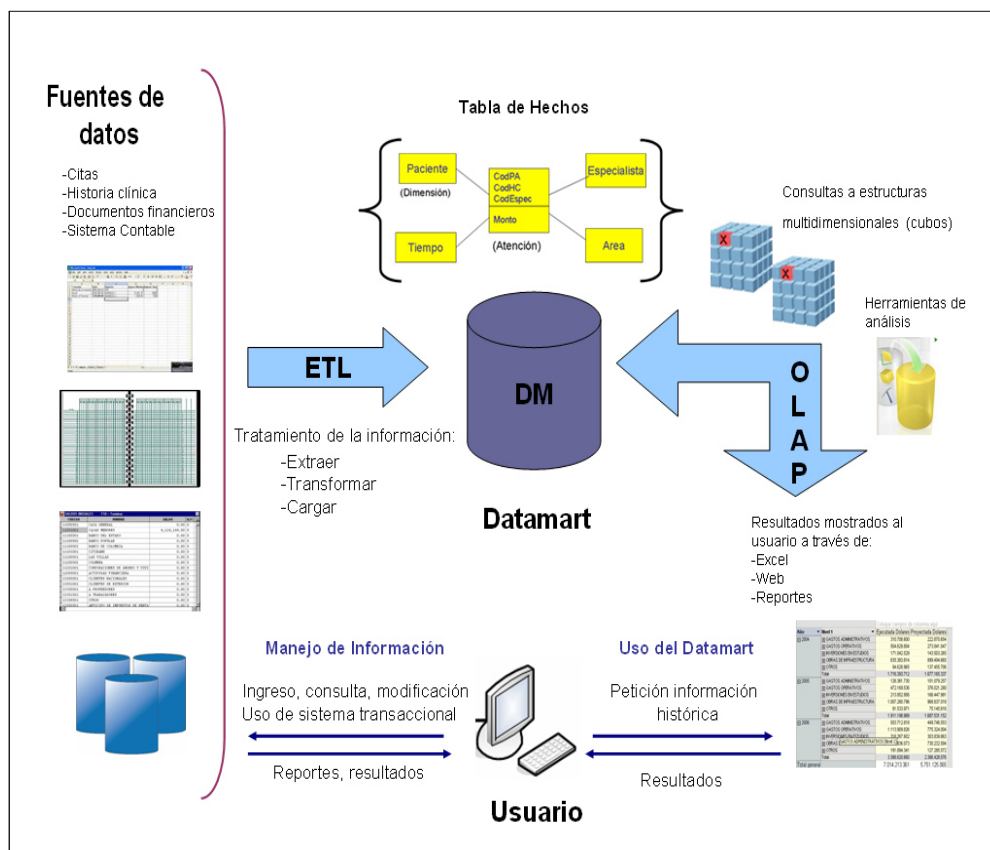


Figura 3.2 Esquema de Data Mart para un Consultorio.

## CASO 2: UNA METODOLOGÍA PARA SECTORIZAR PACIENTES EN EL CONSUMO DE MEDICAMENTOS APLICANDO DATAMART Y DATAMINING EN UN HOSPITAL NACIONAL

Este trabajo, propone un método para el análisis de datos, para evaluar la forma con la que se consumen los medicamentos en un hospital peruano, poder identificar algunas realidades o características no observables que producirían desabastecimiento o insatisfacción del paciente, y para que sirva como una herramienta en la toma de decisión sobre el abastecimiento de medicamentos en el hospital.

El Data mart desarrollado con el nuevo esquema dimensional se alimenta de los datos residentes en una base de datos relacional, luego hace la carga desde los archivos planos de los sistemas transaccionales existentes del hospital, por medio de diversos procesos de extracción, transformación y carga, hacia el nuevo repositorio el cual sirve como fuente de datos (Ver Figura 3.3), al cual se le aplica un algoritmo de minería de datos la clasificación e interpretación de resultados.

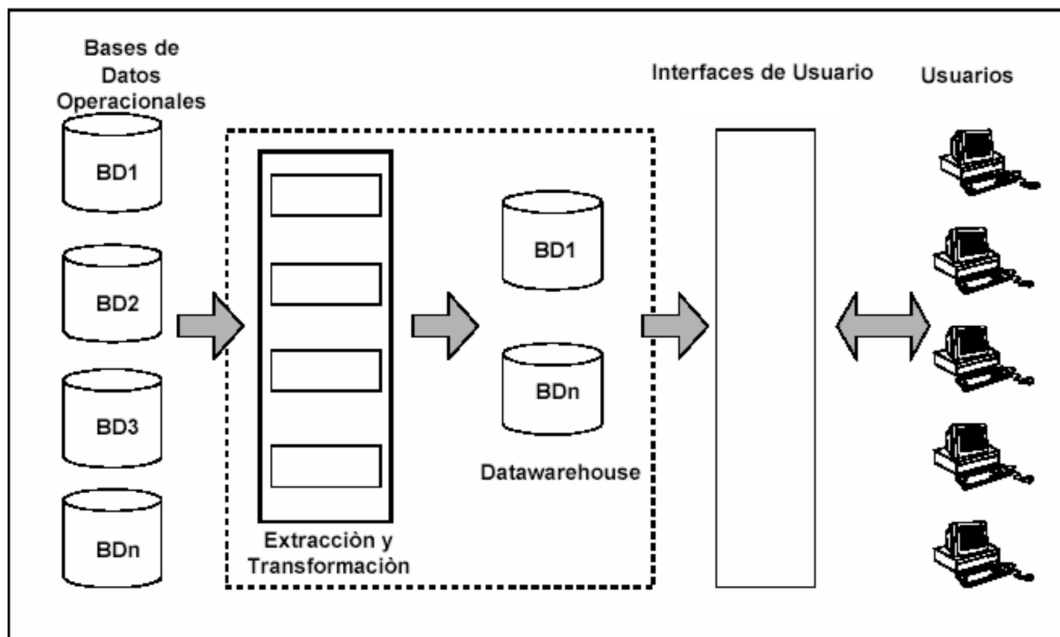


Figura 3.3: Pasos de la Metodología

### **3.3.1. ESTUDIO DE CASOS:**

En ambos casos el Esquema Estrella fue usado porque maneja bien la performance de consultas y reportes que incluyen años de datos históricos y por su simplicidad en comparación con una base de datos normalizada.

A diferencia del Esquema Copo de Nieve, cuyas ventajas de esta normalización son la reducción del tamaño y redundancia en las tablas de dimensión, y un aumento de flexibilidad en la definición de dimensiones, sin embargo, el incremento en la cantidad de tablas hace que se necesiten más operaciones de unión para responder a las consultas, lo que hace lenta la performance, mayor dificultad del entendimiento del modelo y además del mantenimiento que requieren las tablas adicionales.

Existen dos grandes filosofías con respecto a la relación entre los conceptos de Data Warehouse y Data Mart.

Bill Inmon, propuso la idea de que los Data Marts se sirven del Data Warehouse para extraer información. La misma está almacenada en tercera forma normal, en un modelo relacional.

Por su parte, Ralph Kimball, el principal propulsor del enfoque dimensional para el diseño del Data Warehouse, sostiene que el Data Warehouse es el resultado de la unión de los Data Marts de la empresa.

Y en el análisis de los casos, ambos utilizaron el modelo de Kimball donde el ágil desarrollo del Data Mart permitía la obtención de datos concretos además su alcance es más especializado, su análisis y diseño suele tardar mucho menos que en el caso de los Data Warehouses, y por la similitud de la metodología con la del desarrollo del presente trabajo también se utilizará el modelo Kimball.

Pero la desventaja de este modelo es que no resulta siempre sencillo integrar todo en un solo Data Warehouse, debido a que la existencia de múltiples Data Marts diseñada y construida de manera independiente produce inconsistencias y duplicación de datos.

## **CAPÍTULO 4: RESOLUCIÓN DEL PROBLEMA APLICANDO LA TÉCNICA SELECCIONADA**

### **4.1 PROPUESTA METODOLÓGICA.**

Se trata de una investigación que parte de los conceptos fundamentales sobre el desarrollo de aplicaciones de Business Intelligence y que busca adaptar esta base teórica y conceptual a los requerimientos de manejo de información de interés para médicos oncólogos para tomar como referencia los resultados de evaluación del paciente.

El estudio responde a fines prácticos y concretos. Se busca contribuir a mejorar la obtención de información digital obtenida de las Historias Clínicas de los pacientes del Instituto Nacional de Enfermedades Neoplásicas.

La propuesta de implementar un Data Mart en el Departamento de Abdomen del Instituto Nacional de Enfermedades Neoplásicas se fundamenta en las siguientes razones:

- Digitalizar la información del paciente.
- Mejor explotación de la información.
- Información disponibles para los médicos oncólogos.
- Manejo de la información en orden cronológico y por niveles.



**Fuentes de datos**

- CD
- Archivos de Pronunciamientos (html, pdf)
- Revistas / Boletines
- Base de datos

**Tabla de Hechos Factless**

Intensidad	Codigo	Codigos	Fecha	Tiempo
(Dimensiones)	(Expediente)			

**Consultas a estructuras multidimensionales (cubos)**

**ETL** (Extracción, Transformación, Carga) → **DM** (Datamart)

**Tratamiento de la información:**

- Extraer
- Transformar
- Cargar

**OLAP** (Operaciones Analíticas) → **DM**

**Herramientas de análisis**

**Usuario** → **Petición de información histórica** → **Resultados mostrados al usuario**

**Figura 4.1 Esquema del Data Mart.**

## 4.2 IMPLEMENTACIÓN

A continuación, se detallan los pasos que se siguieron en esta metodología:

### 4.2.1 Diseñar el Esquema Dimensional del Data Mart

- **Definición de Dimensiones y Jerarquías.**

Las dimensiones serían las sgtes, como se muestra en la figura 4.2.

Dimensiones:

1. Dimensión Tiempo:
2. Dimensión Complicación Post-Operatoria:
3. Dimensión Médico:
4. Dimensión Paciente:
5. Dimensión Patologías:

A continuación, con sus jerarquías respectivas:

- **Dimensión Persona**
  - Nombre Persona
- **Dimensión Tiempo**
  - Año
  - Mes
- **Dimensión Complicación Post- Operatoria**
  - Complicación
- **Dimensión Médico**
  - Nombre Médico
- **Dimensión Paciente**
  - Nombre Paciente
- **Dimensión Patologías**
  - Diagnóstico

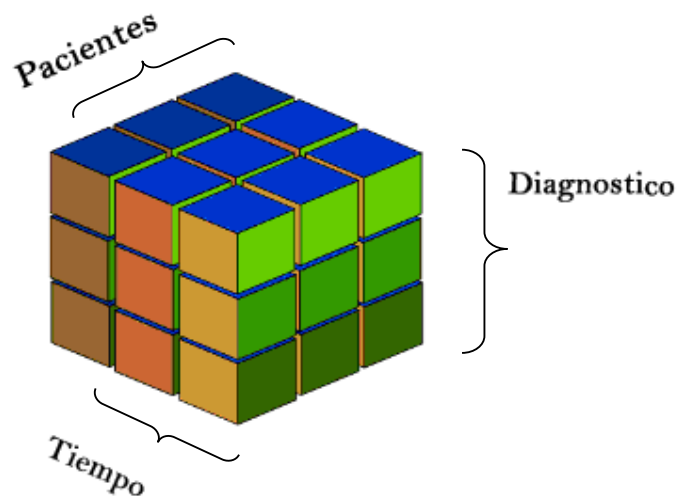


Figura 4.2 Ejemplo de dimensiones usadas en cubo.

- **Definición de la Tabla de Hechos**

### **Creación de las Claves**

Debemos ahora asignar Llaves primarias de cada dimensión, como se muestra en la tabla 4.2:

<b>DIMENSIÓN</b>	<b>CLAVE</b>	<b>TIPO DATO</b>
<b>TIEMPO</b>	<b>Código</b>	<b>Entero</b>
<b>COMPLICACIÓN POST-OPERATORIA</b>	<b>Código</b>	<b>Entero</b>
<b>PATOLOGÍAS</b>	<b>Código</b>	<b>Entero</b>
<b>MÉDICO</b>	<b>Código</b>	<b>Entero</b>
<b>PACIENTE</b>	<b>Código</b>	<b>Entero</b>

Tabla 4.2: Llaves primarias de cada dimensión

### **Creación de Índices**

Por cada dimensión se definieron un índice adicional al que contiene producto de sus claves, generalmente se crearon índices para los nombres o descripciones:

- Dimensión Tiempo: Id\_Tiempo
- Dimensión Complicación Post-Operatoria: Id\_ComplicacionPO
- Dimensión Médico: Id\_Medico
- Dimensión Paciente: Id\_Paciente
- Dimensión Patologías: Id\_Patologia
- Tabla de Hecho: Id\_Morbilidad,  
Id\_Mortalidad,  
Id\_Tiempo,  
Id\_ComplicacionPO,  
Id\_Medico,  
Id\_Paciente,  
Id\_Patologia.

### 4.3 TABLA DE HECHO

En una Tabla de Hechos no se tienen medidas, sólo dimensiones. Este tipo de soluciones es usado en Datamart como de centros de estudios o Centros de Salud, en donde sólo se necesitan conteo.

Por ejemplo en un Hospital, como medida se usaría “pacientes” y luego un contador.

Se caracteriza por:

- No tiene columnas de hechos numéricas.
- Se utilizan para capturar relaciones entre dimensiones.
- Incluyen una columna de hechos ficticia con valor 1 (siempre).

El diseño de nuestra tabla de Hechos tomaría el siguiente modelo. (Figura 4.3)

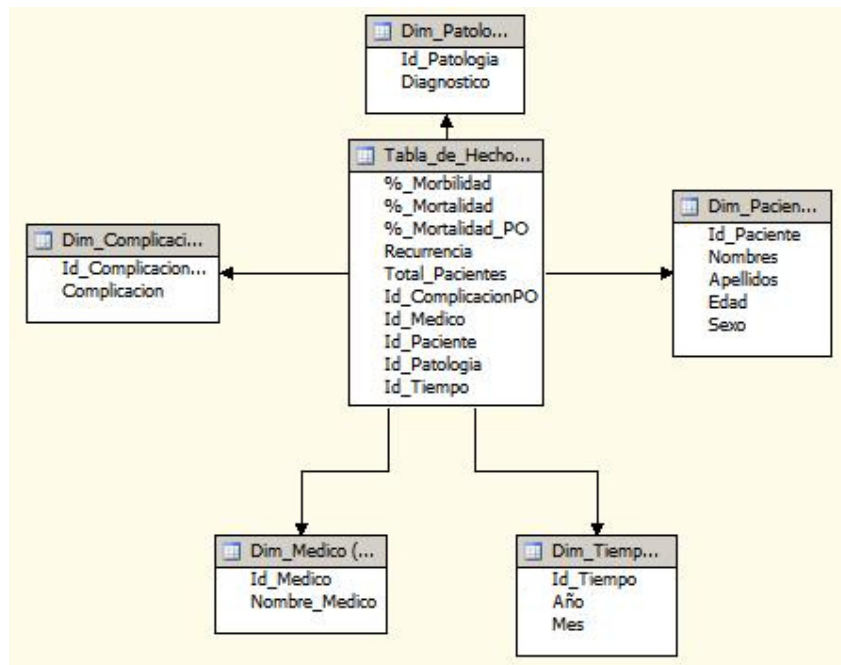


Figura 4.3 Tabla de Hecho.

#### 4.4 DEFINICIÓN DE INDICADORES HOSPITALARIOS

##### Indicador

Variable que es susceptible de ser observada y en lo posible medida, que permite identificar y comparar el nivel o estado de una situación determinada.

Por sus características se han clasificado los Indicadores en:

- Indicadores de Apertura de Historias Clínicas
- Indicadores de Consulta Externa
- Indicadores de Laboratorio
- Indicadores de Aplicaciones, Cirugía Menor, Endoscopias y Procedimientos Especiales
- Indicadores de Sala de Operaciones
- Indicadores de Hospitalización
- Indicadores de Calidad

#### 4.5 INDICADORES DE CALIDAD

##### 4.5.1 Porcentaje de Morbilidad

$$= \frac{\text{Número de Pacientes con determinada Morbilidad}}{\text{Número de Egresos}} \times 100$$

Evalúa la relación existente entre los egresos fallecidos y el total de egresos.

##### 4.5.2 Porcentaje de Mortalidad

$$= \frac{\text{Número de Fallecidos después de 48 Horas}}{\text{Número de Egresos (Vivos y Fallecidos) después de 48 Horas}} \times 100$$

Evalúa la relación existente entre los pacientes fallecidos en el hospital, y los egresos de aquellos pacientes que permanecieron más de 48 horas en el hospital.

##### 4.5.3 Porcentaje de Mortalidad Post-Operatoria

Mide la relación existente entre las muertes producidas dentro de los 30 días siguientes a una intervención quirúrgica.

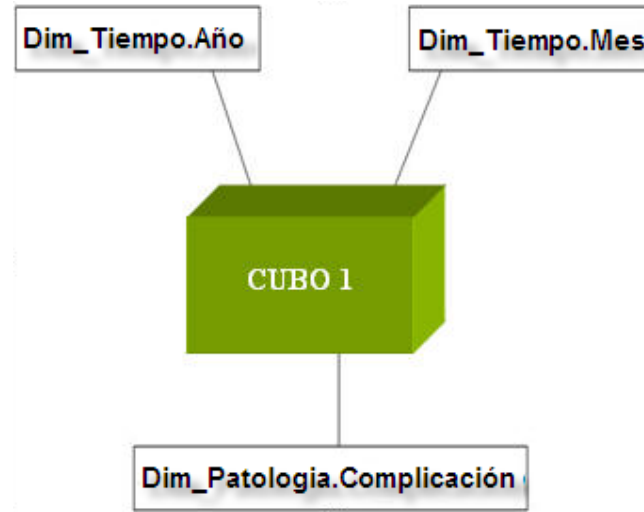
$$= \frac{\text{Número de Fallecidos dentro de los 30 días siguientes a una intervención quirúrgica}}{\text{Número de Intervenciones Quirúrgicas}} \times 100$$

#### 4.6 CUBOS DE INFORMACIÓN

Se ha representado gráficamente 3 cubos de información con sus respectivas dimensiones y jerarquías.

##### **Cubo 1:**

Contiene la cantidad de los pacientes que tuvieron un determinado diagnóstico con su morbilidad y mortalidad post – operatoria asociada, mostrada por año y éste a su vez subdividido en mes (figura 4.4).



**Figura 4.4 Cubo 1 del prototipo.**

Las Dimensiones usadas en este cubo son:

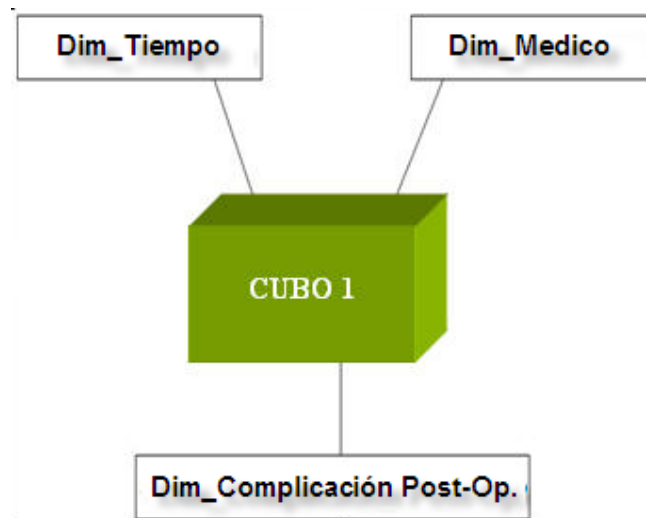
- Dimensión Paciente.
- Dimensión Tiempo.
- Dimensión Patología.

Las jerarquías de la Dimensión Tiempo son:

- Año.
- Mes.

### **Cubo 2:**

Contiene la información del porcentaje de mortalidad y morbilidad post-operatoria de cáncer de páncreas, mostrado por año, mes y muestra información del médico encargado de la cirugía. (Figura 4.5).



**Figura 4.5 Cubo 2 del prototipo.**

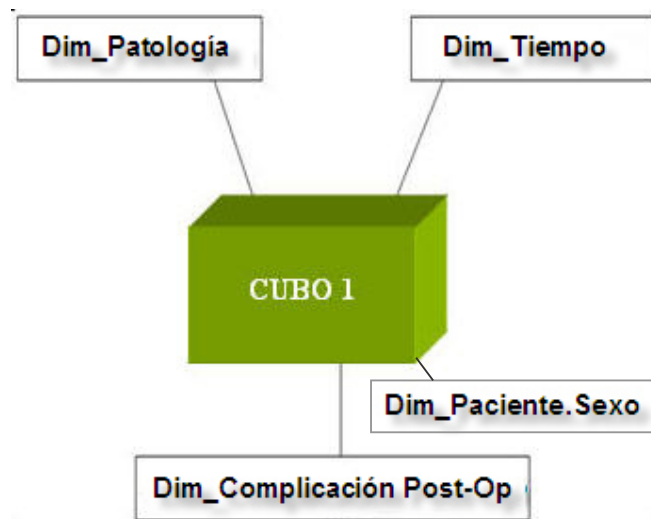
Las Dimensiones usadas en este cubo son:

- Dimensión Médico.
- Dimensión Tiempo.
- Dimensión Complicación Post-Operatoria.

### **Cubo 3:**

- ¿Qué diagnóstico de cáncer con determinada complicación post-operatoria han surgido en los últimos 5 años?

La representación gráfica de este cubo nos muestra la cantidad de pacientes con determinadas complicaciones post-operatorias, diagnóstico de cáncer de páncreas, mostrado por años y sexo. (Figura 4.6).



**Figura 4.6 Cubo 3 del prototipo.**

Las Dimensiones usadas en este cubo son:

- Dimensión Complicación Post – Operatoria.
- Dimensión Tiempo.
- Dimensión Patología.
- Dimensión Paciente.



## **CAPÍTULO 5: DESCRIPCIÓN DE LA SOLUCIÓN TECNOLÓGICA**

### **5.1 HERRAMIENTAS UTILIZADAS**

#### **5.1.1 Introducción:**

La elección de las herramientas utilizadas, se basó en la facilidad de utilización, adquisición de las licencias obtenidas por la institución y el entrenamiento del personal en el uso de estas herramientas.

#### **5.1.2 ETL (Extraer, Transformar y Cargar)**

La información fuente se encuentra en una base de datos SQL 2005 Server.

El proceso de ETL (Extraer, Transformar y Cargar) del Data Mart consistirá en ingresar la información de estos archivos a la BD, para las pruebas solo se tomara una muestra referencial.

#### **Herramientas para implementar procesos ETL:**

SQL Server 2005 ofrece múltiples herramientas para la implementación de los procesos ETL:

- Transact – SQL.
- Consultas Distribuidas.
- Utilitario BCP y la instrucción BULK INSERT
- SQL Server Integration Services (SSIS)

Al tener la información en una base de datos SLQ 2005, se identifico la información relevante debería ser extraída mediante sentencias Transact – SQL, ya que mediante procedimientos almacenados nos permitiría ingresar la información elegida como campos al esquema deseado puesto que nuestras tablas provienen de una sola fuente (SQL Server 2005) y sería cargada a otra base de datos SQL Server 2005, mejorando así la performance y aminorando el flujo de carga.

Las demás herramientas de implementación de procesos ETL permiten cargar datos desde otras bases de datos y diferentes fuentes de datos, lo cual no nos facilitaría el proceso ETL ya que en nuestro caso solo estamos realizando el proceso ETL dentro de nuestra Base de Datos SQL 2005 y de un solo tipo de origen de datos.

### **5.1.3 Microsoft SQL Server Management Studio 2005**

*Microsoft SQL Server Management Studio* es un entorno integrado para obtener acceso a todos los componentes de SQL Server, así como para configurarlos y administrarlos. *Microsoft SQL Server Management Studio* combina un amplio grupo de herramientas gráficas con un editor de texto enriquecido para ofrecer acceso a SQL Server a los programadores y administradores, sin importar su nivel de especialización.

*Microsoft SQL Server Management Studio* combina las funciones del Administrador corporativo y el Analizador de consultas, herramientas incluidas en versiones anteriores de SQL Server, en un único entorno. Además, *Microsoft SQL Server Management Studio* proporciona un entorno para administrar *Analysis Services*, *Integration Services*, *Reporting Services* y *XQuery*. Este entorno ofrece a los programadores una experiencia familiar y proporciona a los administradores de bases de datos una herramienta única para realizar sus tareas con la facilidad de las herramientas gráficas y una experiencia de secuencias de comandos enriquecida.

Utilizando *Microsoft SQL Server Management Studio 2005*, se construyó nuestra Base de Datos Datamart\_INEN, en el cual se cargó (utilizando sentencias Transact-Sql) nuestras dimensiones y la Tabla de Hechos se cargó de datos utilizando las fórmulas de los indicadores de calidad según los campos requeridos, como se muestra en la figura 5.1.

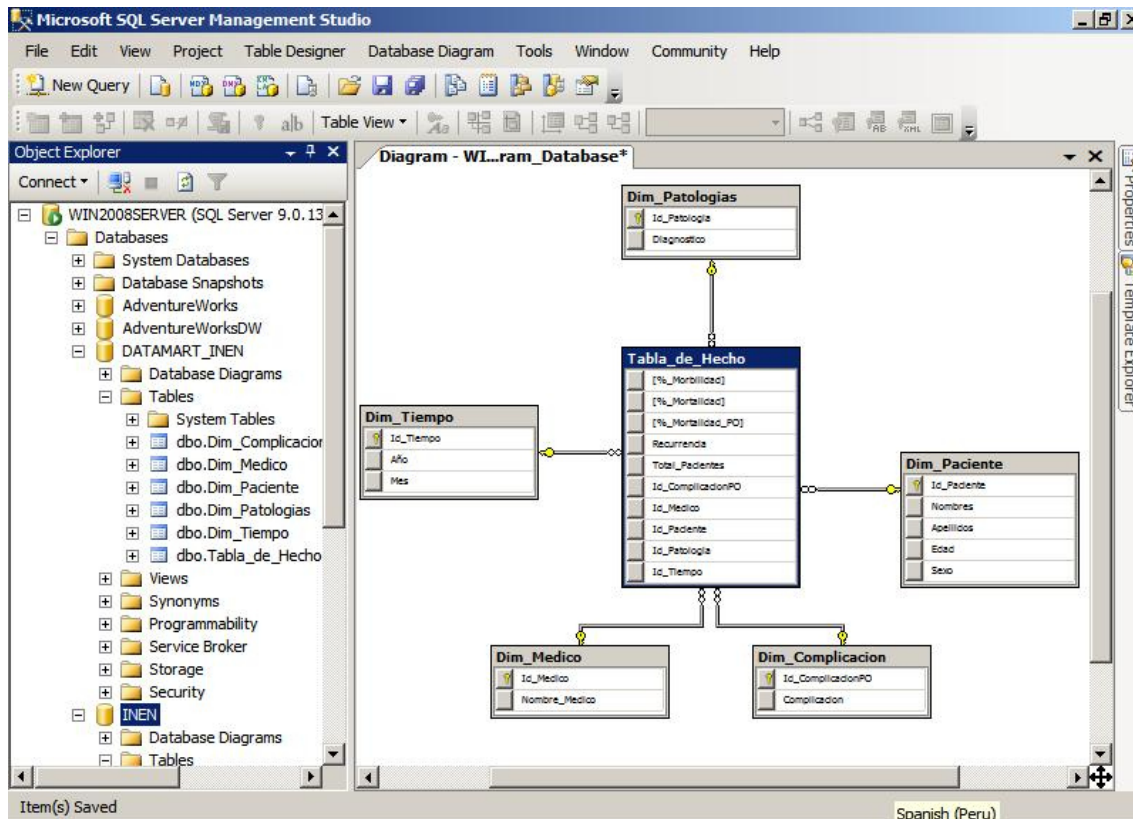


Figura 5.1: Desarrollo de la BD en SQL Server 2005 para el uso del Data mart.

#### 5.1.4 Business Intelligence Development Studio

*SQL Server Business Intelligence Development Studio* es un entorno integrado para desarrollar construcciones de inteligencia empresarial, como cubos, orígenes de datos, informes y paquetes de *Integration Services*. *Business Intelligence Development Studio* incluye plantillas de proyecto que proporcionan un contexto para desarrollar construcciones específicas. Por ejemplo, se puede optar por un proyecto de *Analysis Services* si el objetivo es crear una base de datos de *Analysis Services* que contenga cubos, dimensiones o modelos de minería de datos.

En *SQL Server Business Intelligence Development Studio*, es posible desarrollar proyectos que formen parte de una solución independiente de un servidor concreto. Por ejemplo, puede incluir un proyecto de *Analysis Services*, de *Integration Services* y de *Reporting Services* en la misma solución. Puede implementar los objetos en un servidor de prueba para probarlos durante el desarrollo y posteriormente,

implementar el resultado de los proyectos en uno o más servidores de ensayo o de producción.

En *SQL Server Business Intelligence Development Studio* se realizó la conexión a nuestra base de datos *Datamart\_INEN* en *Microsoft SQL Server Management Studio*, cuyo diagrama de la base de datos se muestra en la figura 5.2.

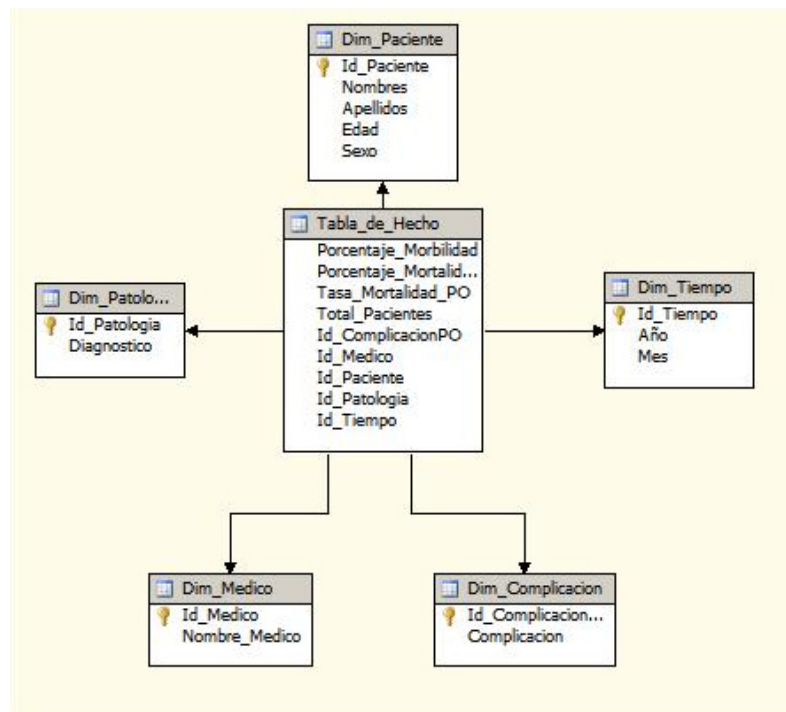


Figura 5.2: Diagrama de la Base de Datos en *Microsoft SQL Server Management Studio*.

Luego se realizó la carga de la vista de orígenes de datos (*Data source views*) desde nuestro origen de datos (*Data source*); los *Data source views* almacenan la estructura de las tablas y sus relaciones entre ellas, el cual nos permite almacenar de forma persistente en nuestro proyecto de análisis, por lo tanto, nos permite trabajar sin necesidad de una conexión permanente hacia nuestra base de datos de origen (*Datamart\_INEN*).

Posteriormente se realizó la creación de nuestros cubos de información, como se muestra en la figura 5.3.

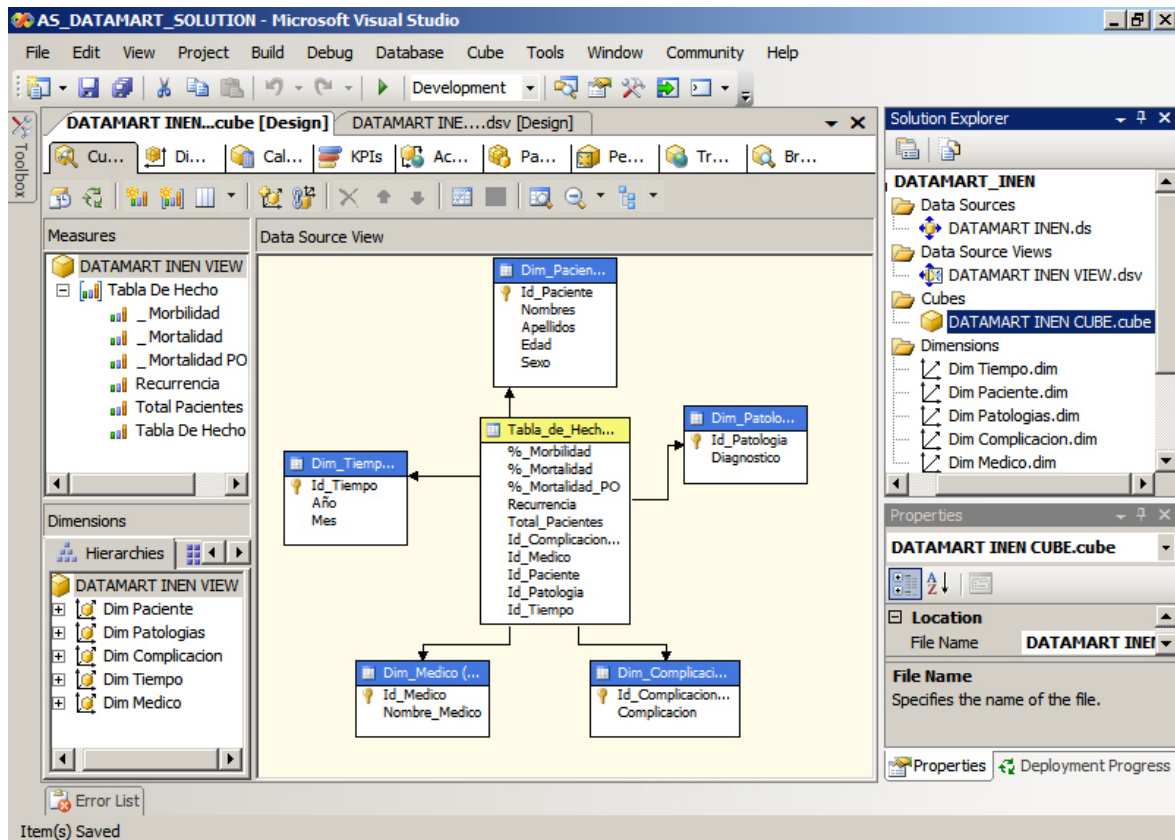


Figura 5.3: Desarrollo de la solución BI en *Business Intelligence Development Studio*.

### 5.1.5 Informes del Data Mart

Se realizaron pruebas con los informes usando Microsoft Office Excel 2007, que es un entorno en el que se pueden crear diversos tipos de informes a partir de distintos orígenes de datos, en nuestro caso Microsoft Excel 2007 ya no necesita saber los nombres de servidor o de base de datos de los orígenes de datos corporativos. En su lugar, puede utilizar el inicio rápido para seleccionar de una lista de orígenes de datos que el administrador o el experto del grupo de trabajo haya puesto a su disposición. Un administrador de conexiones de Excel le permite ver todas las conexiones y facilita la tarea de volver a utilizar una conexión o de sustituir una conexión por otra. Permitiendo crear informes altamente personalizables y modelos analíticos que se pueden actualizar automáticamente para incluir los datos del Data Mart, y además ésta herramienta ofrece una rápida conversión en gráficos e impresión de dichos informes.

## CAPÍTULO 6: ANÁLISIS E INTERPRETACIÓN DE RESULTADOS

Para poder realizar las pruebas de funcionamiento del prototipo del Data Mart se tomó una muestra de información de los años 2002 – 2007 de las historias clínicas de los pacientes con cáncer de páncreas en el Instituto Nacional de Enfermedades Neoplásicas, llenando esta información en nuestra Base de Datos en SQL Server 2005 y luego construyendo nuestro modelo multidimensional para poder llenarlos con las sentencias Transact – SQL.

En el análisis del primer cubo que contiene la cantidad de los pacientes que tuvieron un determinado diagnóstico con su morbilidad y mortalidad post – operatoria asociada, en la figura 6.1 se muestra por año y este a su vez subdividido en mes, donde se aprecia que la cantidad total de pacientes es 60, y se tiene un alto porcentaje de morbilidad de 29.88.

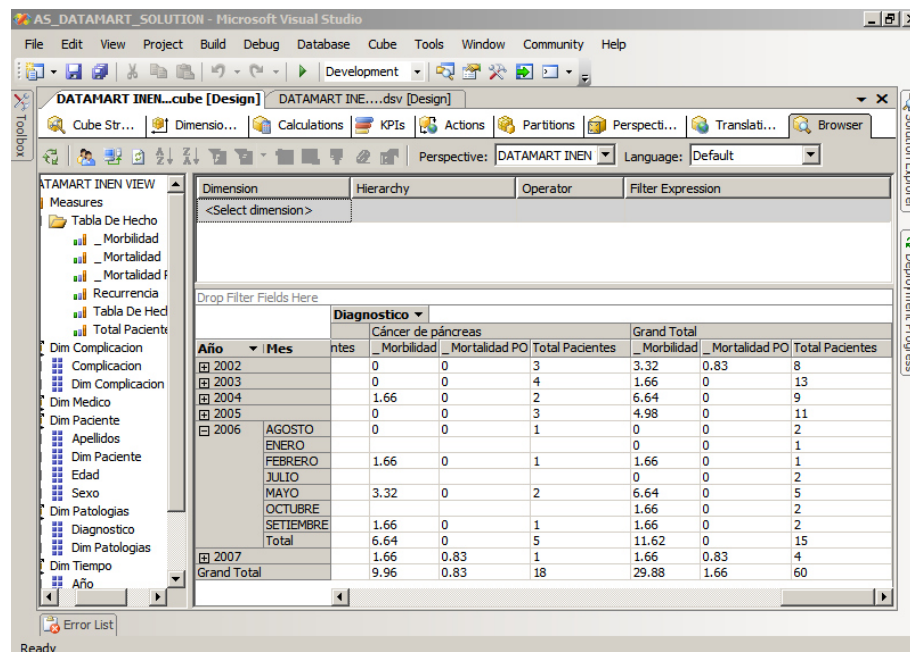


Figura 6.1: Resultados de la Elaboración del Cubo1

En la Figura 6.2 se exporto el cubo a Excel 2007, donde se puede apreciar que en los últimos 5 años (2002 - 2007), la mayor incidencia de morbilidad post – operatoria es dada al cáncer de ampolla de váter con un porcentaje de 14.94, luego seguida de cáncer de páncreas con un porcentaje de 9.96 y con menor porcentaje (1.66) se encuentra el cáncer de duodeno, el cáncer periampular de cabeza de páncreas y el carcinoma solido.

También se aprecia que la mortalidad post-operatoria es encontrada en baja proporción en los pacientes estudiados con cáncer de páncreas.

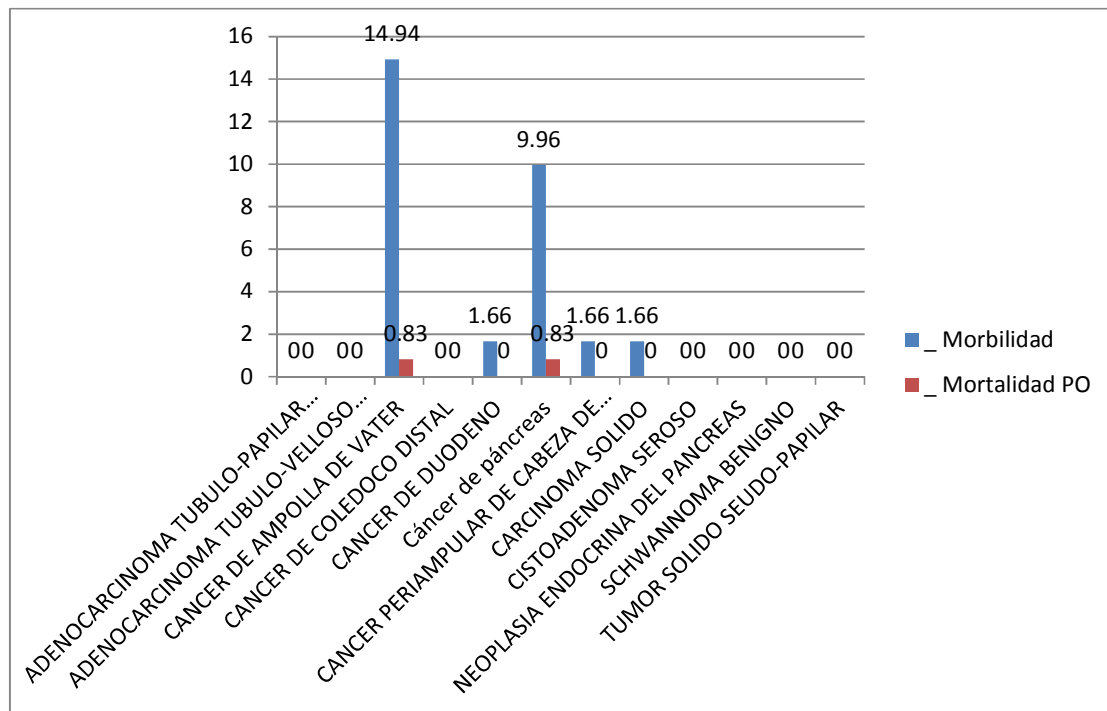


Figura 6.2: Representación gráfica del Cubo 1

En el cubo 2 del prototipo (Figura 6.3), se puede apreciar al médico oncólogo y las complicaciones aparecidas en sus pacientes operados, como así también se muestra los años y meses que ocurrieron dichas operaciones.

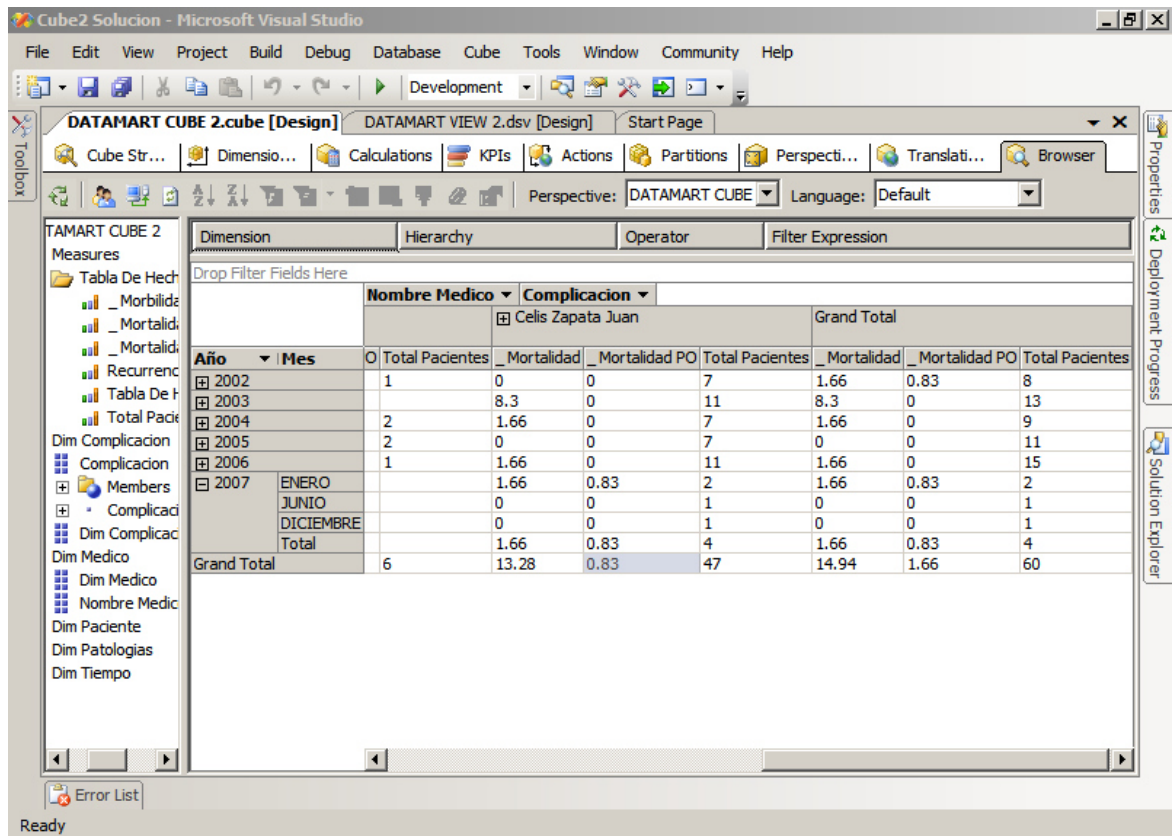


Figura 6.3: Representación gráfica del Cubo 2

En la tabla 6.1 muestra las estadísticas de los médicos oncólogos del periodo 2002 – 2007 en el Instituto Nacional de Enfermedades Neoplásicas, la cantidad de pacientes que operaron y sus porcentajes de morbilidad y mortalidad post – operatoria asociados.

Médicos	Valores		
	% Morbilidad	% Mortalidad Post_Oper	Total Pacientes
Medico 1	4.98	0	6
Medico 2	18.26	0.83	47
Medico 3	1.66	0	1
Medico 4	4.98	0.83	6
<b>Total general</b>	<b>29.88</b>	<b>1.66</b>	<b>60</b>

Tabla 6.1: Información de los Médicos y valores asociados.



En la figura 6.4 muestra gráficamente que el médico 2 ha tenido una mayor cantidad de pacientes operados y a la vez el mayor porcentaje de pacientes con una alta tasa de morbilidad, en el cual se podría sugerir un seguimiento a los médicos con un alto porcentaje de morbilidad y mortalidad post-operatoria con el fin de poder mejorar las técnicas usadas por los médicos en la operación o hacer un seguimiento a las condiciones pre-operatorias del paciente con el fin de reducir esta estadística todavía considerablemente alta.

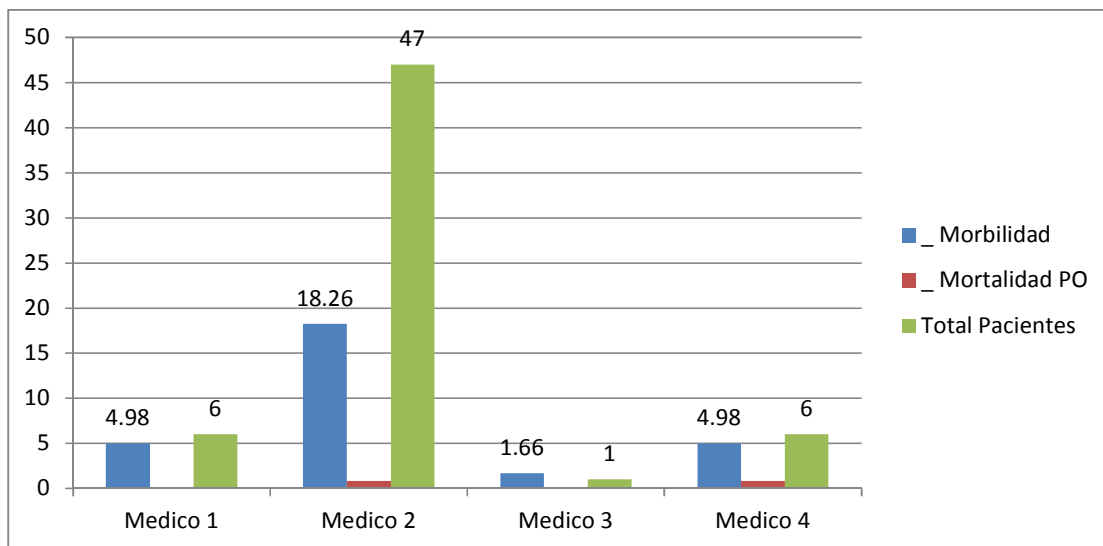


Figura 6.4: Representación gráfica de resultados del Cubo 2

En el cubo 3 se puede apreciar que la representación gráfica nos muestra la cantidad de pacientes con determinadas complicaciones post-operatorias y su diagnóstico de cáncer de páncreas asociado, agrupado por años y sexo del paciente como se muestra en la figura 6.5.

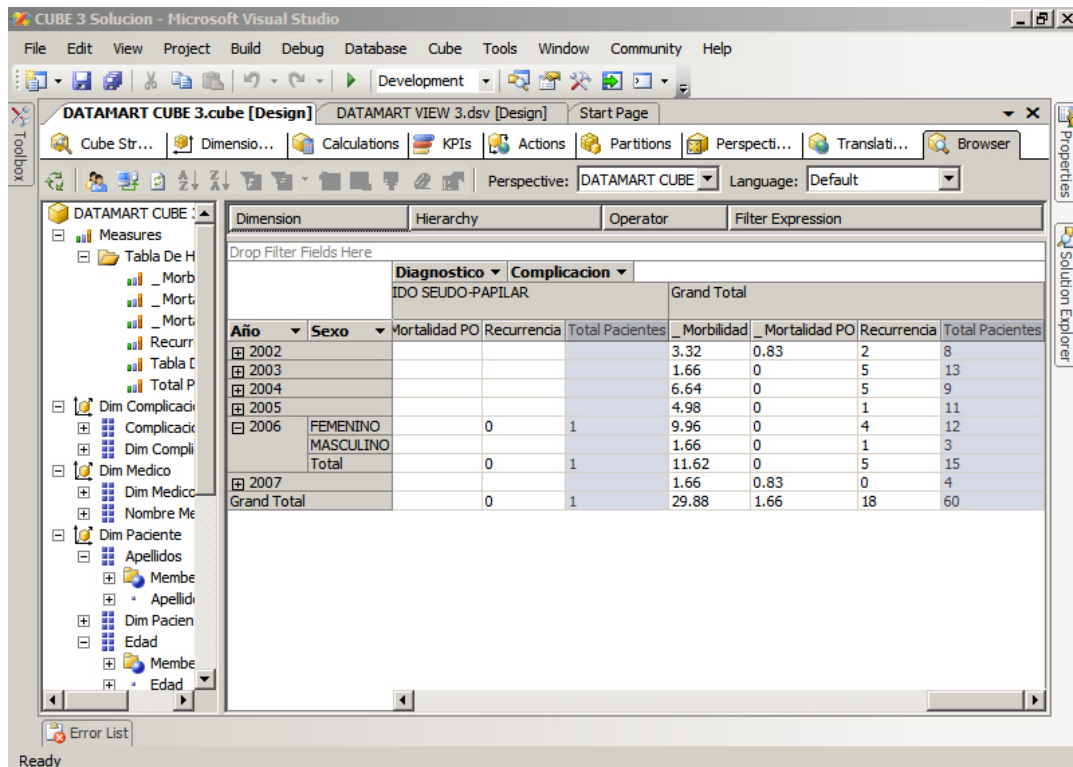


Figura 6.5: Representación gráfica del cubo 3

Esta relación nos ayudaría a encontrar que complicaciones post-operatorias tendrían una mayor relación con un determinado diagnóstico de cáncer de páncreas, en Tabla 6.2 encontramos que la fístula pancreática tiene mayor incidencia en pacientes con Cáncer de Ampolla de Váter, por lo cual podemos realizar un trabajo de investigación con mayor profundidad para encontrar más variables relacionadas con este resultado.

En donde se tomarían en cuenta variables como por ejemplo: los cirujanos con mayor experiencia, la capacidad de realizar una intervención en menos tiempo y con menor pérdida de sangre, mejores cuidados pre y postoperatorios, mejor técnica anestésica y un soporte nutricional adecuado.

<b>Diagnóstico / Complicación Post-Operatoria</b>	<b>%Morbilidad</b>	<b>Total Total Pacientes</b>
<b>COLECCIÓN INTRA ABDOMINAL</b>		<b>1</b>
CANCER DE AMPOLLA DE VATER		1
<b>DEHISCENCIA H.O</b>	<b>1.66</b>	<b>1</b>
CANCER DE AMPOLLA DE VATER	1.66	1
<b>Distress respiratorio</b>		<b>1</b>
CANCER DE AMPOLLA DE VATER		1
<b>FISTULA PANCREATICA</b>	<b>1.66</b>	<b>7</b>
CANCER DE AMPOLLA DE VATER	1.66	4
CANCER DE DUODENO		1
Cáncer de páncreas		1
CARCINOMA SOLIDO		1
<b>HEMORRAGIA POST OPERATORIA INTRA ABDOMINAL</b>		<b>3</b>
Cáncer de páncreas		3
<b>HERPES ZOSTER</b>		<b>1</b>
Cáncer de páncreas		1
<b>PROCTITIS CRONICA</b>		<b>2</b>
CANCER DE AMPOLLA DE VATER		1
Cáncer de páncreas		1
<b>Retardo Vaciamiento Gástrico</b>		<b>4</b>
CANCER DE AMPOLLA DE VATER		2
Cáncer de páncreas		1
CANCER PERIAMPULAR DE CABEZA DE PANCREAS		1
<b>Sin Complicación</b>	<b>0</b>	<b>40</b>
ADENOCARCINOMA TUBULO-PAPILAR PERIAMPULAR CON EXTENSION EN CABEZA DE PANCREAS		1
ADENOCARCINOMA TUBULO-VELLOSO DE LA AMPOLLA DE VATER		1
CANCER DE AMPOLLA DE VATER	0	12
CANCER DE COLEDOCO DISTAL	0	8
CANCER DE DUODENO		3
Cáncer de páncreas	0	11
CISTOADENOMA SEROSO		1
NEOPLASIA ENDOCRINA DEL PANCREAS		1
SCHWANNOMA BENIGNO		1
TUMOR SOLIDO SEUDO-PAPILAR		1
<b>Total general</b>	<b>3.32</b>	<b>60</b>

Tabla 6.2: Cantidad de pacientes con determinadas complicaciones post-operatorias.

## **CAPÍTULO 7: CONCLUSIONES Y FUTUROS TRABAJOS**

Es importante conocer informes estadísticos de las enfermedades neoplásicas a nivel nacional, para reconocer e investigar de manera integral las enfermedades de mayor frecuencia por cada región, departamento o ciudad, más aún contando con la información de años anteriores.

El cáncer de páncreas está incluido entre las 5 primeras causas de mortalidad, de etiología en estudio, pero es interesante se investigue sobre su prevalencia en diferentes entidades y ciudades, para reconfirmar que es una neoplasia de importante frecuencia en nuestro medio, o está ascendiendo, como en otros países se está en evaluación.

Pese a los inconvenientes que puedan tenerse en los sistemas de registro ya que todavía se realiza en papel físico, la información presentada podrá servir para la toma de decisiones con lo que ocurre con la enfermedades neoplásicas digestivas y demás neoplasias del cuerpo, respecto a su frecuencia de presentación y su distribución en los casos presentados, siendo imprescindible investigar de manera integral la etiología, prevalencia, incidencia y factores de riesgo de este importante grupo de patologías, de modo que basados en evidencias se puedan orientar mejor las medidas para su prevención y control en nuestra población.

El presente trabajo se puede aplicar tanto a los demás departamentos del Instituto Nacional de Enfermedades Neoplásicas como a los Institutos Regionales de Enfermedades Neoplásicas (Trujillo y Arequipa), pudiendo así todos los Departamentos de Cáncer (cabeza y cuello, tórax, neuro-oncología, urología, etc.) utilizar estas herramientas y no sólo realizar toma de decisiones con mayor prontitud, sino que también como medida de control e investigación sobre las demás variables implicadas en el estado pre y post – operatorio del

paciente con cáncer. Llegando a cumplir con mayor eficiencia las metas sobre las medidas de los índices de calidad establecidos en el INEN.

Y como medida de mejora se puede implementar mecanismos gráficos para mostrar la información consolidada como el Dashboard o Tablero de control, donde nos mostraría la información estratégica requerida para la toma de decisiones.

## ANEXOS

TABLA 1:

**Table I.** Proposed definition of postpancreatectomy hemorrhage (PPH)

<i>Time of onset</i>
- Early hemorrhage ( $\leq 24$ h after the end of the index operation)
- Late hemorrhage ( $> 24$ h after the end of the index operation)
<i>Location</i>
- Intraluminal (intraenteric, eg, anastomotic suture line at stomach or duodenum, or pancreatic surface at anastomosis, stress ulcer, pseudoaneurysm)
- Extraluminal (extraenteric, bleeding into the abdominal cavity, eg, from arterial or venous vessels, diffuse bleeding from resection area, anastomosis suture lines, pseudoaneurysm)
<i>Severity of Hemorrhage</i>
<i>Mild</i>
- Small or medium volume blood loss (from drains, nasogastric tube, or on ultrasonography, decrease in hemoglobin concentration $< 3$ g/dl)
- Mild clinical impairment of the patient, no therapeutic consequence, or at most the need for noninvasive treatment with volume resuscitation or blood transfusions (2-3 units packed cells within 24 h of end of operation or 1-3 units if later than 24 h after operation)
- No need for reoperation or interventional angiographic embolization; endoscopic treatment of anastomotic bleeding may occur provided the other conditions apply
<i>Severe</i>
- Large volume blood loss (drop of hemoglobin level by $\geq 3$ g/dl)
- Clinically significant impairment (eg, tachycardia, hypotension, oliguria, hypovolemic shock), need for blood transfusion ( $> 3$ units packed cells)
- Need for invasive treatment (interventional angiographic embolization, or relaparotomy)

**Table II.** Proposed classification of PPH: clinical condition, diagnostic and therapeutic consequences

<i>Grade</i>	<i>Time of onset, location, severity and clinical impact of bleeding</i>		<i>Clinical condition</i>	<i>Diagnostic consequence</i>	<i>Therapeutic consequence</i>
A	Early, intra- or extraluminal, mild		Well	Observation, blood count, ultrasonography and, if necessary, computed tomography	No
B	Early, intra- or extraluminal, severe	Late, intra- or extraluminal, mild*	Often well/intermediate, very rarely life-threatening	Observation, blood count, ultrasonography, computed tomography, angiography, endoscopy†	Transfusion of fluid/blood, intermediate care unit (or ICU), therapeutic endoscopy,† embolization, relaparotomy for early PPH
C		Late, intra- or extraluminal, severe	Severely impaired, life-threatening	Angiography, computed tomography, endoscopy†	Localization of bleeding, angiography and embolization, (endoscopy†) or relaparotomy, ICU

ICU, Intensive care unit; PPH, postpancreatectomy hemorrhage.

\*Late, intra- or extraluminal, mild bleeding may not be immediately life threatening to patient but may be a warning sign for later severe hemorrhage ("sentinel bleed") and is therefore Grade B.

†Endoscopy should be performed when signs of intraluminal bleeding are present (melena, hematemesis, or blood loss via nasogastric tube).

TABLA 2:

**Table II.** Main parameters for POPF grading

<i>Grade</i>	<i>A</i>	<i>B</i>	<i>C</i>
Clinical conditions	Well	Often well	Ill appearing/ bad
Specific treatment*	No	Yes/no	Yes
US/CT (if obtained)	Negative	Negative/ positive	Positive
Persistent drainage (after 3 weeks)†	No	Usually yes	Yes
Reoperation	No	No	Yes
Death related to POPF	No	No	Possibly yes
Signs of infections	No	Yes	Yes
Sepsis	No	No	Yes
Readmission	No	Yes/no	Yes/no

US, Ultrasonography; CT, computed tomographic scan; POPF, postoperative pancreatic fistula.

\*Partial (peripheral) or total parenteral nutrition, antibiotics, enteral nutrition, somatostatin analogue and/or minimal invasive drainage.

†With or without a drain in situ.

TABLA 3:

**Table II.** Consensus definition of DGE after pancreatic surgery

<i>DGE grade</i>	<i>NGT required</i>	<i>Unable to tolerate solid oral intake by POD</i>	<i>Vomiting/gastric distension</i>	<i>Use of prokinetics</i>
A	4–7 days or reinsertion > POD 3	7	±	±
B	8–14 days or reinsertion > POD 7	14	+	+
C	>14 days or reinsertion > POD 14	21	+	+

*DGE*, Delayed gastric emptying; *POD*, Postoperative day; *NGT*, Nasogastric tube.

To exclude mechanical causes of abnormal gastric emptying, the patency of either the gastrojejunostomy or the duodenojejunostomy should be confirmed by endoscopy or upper gastrointestinal gastrographin series.

**Table III.** Parameters for grading of DGE

<i>DGE</i>	<i>Grade A</i>	<i>Grade B</i>	<i>Grade C</i>
Clinical condition	Well	Often well/minor discomfort	Ill/bad/severe discomfort (increased overall risk owing to complications and procedures)
Comorbidities	No	Possibly yes (pancreatic leak or fistula, intraabdominal abscess)	Possibly yes (pancreatic leak or fistula, intraabdominal abscess)
Specific treatment	Possibly yes (prokinetic drugs)	Yes (prokinetic drugs, potential reinsertion of NGT)	Yes (prokinetic drugs, NGT)
Nutritional support (enteral or parenteral)	Possibly yes (slower return to solid food intake)	Yes (partial parenteral nutrition)	Yes (total parenteral or enteral nutrition via NGT, prolonged, i.e., >3 weeks postoperatively)
Diagnostic evaluation	No	Possibly yes (endoscopy, upper GI contrast study, CT)	Yes (endoscopy, upper GI contrast study, CT)
Interventional treatment	No	No	Possibly yes (e.g., abscess drainage, relaparotomy for complication, relaparotomy for DGE)
Prolongation of hospital stay	Possibly yes	Yes	Yes
Delay of potential adjuvant therapy	No	No	Yes

*CT*, Computed tomography; *DGE*, Delayed gastric emptying; *GI*, Gastrointestinal; *NGT*, nasogastric tube.



## REFERENCIAS BIBLIOGRÁFICAS

### Libros:

- [Dorland97] Dorland Diccionario Enciclopédico Ilustrado de Medicina. 28ª ed. Madrid: McGraw-Hill, Interamericana; 1997; p. 537
- [Estremadoyro+96] Estremadoyro O, Alvarez Y, Estremadoyro L & Gamero D. Cáncer de páncreas en el Hospital Regional Honorio Delgado de Arequipa. Rev. Gastroenterol del Perú. 1996; 16:49-61.
- [Inmon05] W. H. Inmon, Building the Data Warehouse (2005) Publisher by Wiley Publishing, Inc.
- [Inen95] Registro de cáncer de Lima metropolitana 1990-1991. Publicación del Centro de Investigación en Cáncer "Maes-Heller". INEN. (1995).
- [Nago94] Nago, A. Epidemiología de las enfermedades digestivas. Enfermedades del aparato digestivo en el Perú. En: Gastroenterología III. Sociedad de Gastroenterología del Perú. Lima - Perú 1994. Pg. 5 -15.
- [Polanco99] Dr. Victor Polanco Cano Cirugía: I cirugía general (1999)
- [RCLM98] Registro de Cáncer de Lima Metropolitana 1990 - 1993. Publicación del Centro de Investigación en cáncer "Maes-Heller". Vol II. Agosto 1998.
- [Rochester75] Universidad de Rochester, Oncología Clínica, Publicación de la Sociedad Americana de Cancer, 5ta Edición, 1980.
- [Solidoro+75] Solidoro Andres, Caceres, Eduardo, De la Flor Jorge, Montalberti Juan, Salem Luis, Instituto Nacional de Enfermedades Neoplásicas, Oncologia, 1975.
- [Solidoro05] Solidoro Santisteban Andrés, Apuntes de Cancerología, Primera Edición, (2005).
- [Wiley02] John Wiley and Sons, The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling (Second Edition) (2002).
- [Zvenger05] Patricia Andra Zvenger. Introducción al soporte de Decisiones. Incorporación de Soluciones OLAP en entornos empresariales. Dpto. Ciencias e Ing. De la Computación. Univ. Nacional del Sur. Diciembre 2005.

**Revistas:**

- [Farfan+00] FARFAN, Gustavo y CABEZAS, César. Mortalidad por enfermedades digestivas y hepatobiliares en el Perú, 1995-2000. *Rev. gastroenterol. Perú*. Oct./Dic. 2002, vol.22, N° 4 [citado 22 Mayo 2008]; p.310-323.
- [Díaz+96] Díaz J. Tantalean A., y Col.; Cáncer primario de vesícula biliar. *Rev.Gastroent.Perú* 1996; 16:142-147.
- [Rivera+98] Fabiola Rivera, Augusto Nago, Alejandro Bussalleu, Ruth Medina, Olga Lizama, G. Loayza. Cáncer gástrico en pacientes menores de 40 años en el Hospital A. Loayza entre 1991 - 1998. *Revista de Gastroenterología del Perú - Volumen 20, N°3, 2000.*
- [Espejo 96] Espejo, H. Cáncer gástrico. *Rev de Gastroenterol del Perú*. 1996; 16:11-12
- [Berrospi+97] Ruiz E, Berrospi F, Morante C, Payet E, Celis J & Montalbetti J. Operabilidad y resecabilidad del cáncer gástrico: análisis de 2880 casos en 15 años. *Rev. Gastroenterol. Del Perú*. 1997; 17:135-142.
- [RMFIHU03] Diagnóstico, *Revista Médica de la Fundación Instituto Hipolito Unanue*. Enero – Febrero 2003 Volumen 42, N° 1; 4-5
- [Anderson06] OncoLog, M. D. Anderson Cancer Center, Julio - Agosto 2006, Vol. 51, No. 7; pág. 5-6.
- [REED04] *Revista Española de Enfermedades Digestivas*; Nov. 2004; Volumen 96 N° 11.

**Tesis:**

- [Agreda+07] Agreda Huamán, Peter Guillermo y Sanchez Mercado María Elena, Implementación de Datamart sobre Jurisprudencia - Tribunal Constitucional JURISDAT, Tesis, (2007), Lima-Perú.
- [Poquioma95] Poquioma, E. Tendencia en la incidencia de cáncer en Lima metropolitana 1968 - 1991. Tesis para optar el Título de Médico Cirujano. UPCH. Lima Perú, (1995).
- [Rivas06] Tapia Rivas, Iván Gildo, Una metodología para sectorizar pacientes en el consumo de medicamentos aplicando Datamart y Datamining en un Hospital Nacional, Tesis, (2006), Lima- Perú.
- [Valencia08] Valencia Mariñas, Hugo, Fístula pancreática en función al control sistemático de amilasa en pacientes sometidos a duodenopancreatectomía en el Instituto Nacional de Enfermedades Neoplásicas 2002-2007, Tesis Maestría, (2008), Lima – Perú.

**Direcciones Electrónicas:**

Instituto Nacional de Enfermedades Neoplásicas, Perfil Epidemiológico, INEN.  
<http://www.inen.sld.pe/intranet/estadepidemiologicos.htm#2>  
22/01/2008

Data Warehousing: Similarities and Differences of Inmon and Kimball.  
<http://www.b-eye-network.com/view/743>

Data Warehousing: Relational vs. Multi-Dimensional Data.  
<http://www.b-eye-network.com/view/757?jsessionid=a5045377d071158c71f14bfb912a1c64>

Planeación de informes de Excel en cubos OLAP.  
[http://technet.microsoft.com/es-es/library/cc263194\(TechNet.10\).aspx](http://technet.microsoft.com/es-es/library/cc263194(TechNet.10).aspx)